ARCO PATENT OFFICE

Bo-eki Bldg., 123 Higashi-machi, Chuo-ku, Kobe 650-0031, Japan
P.O.Box 3, Kobe Port, Kobe 651-0191, Japan
Phone +81-78-321-8822 / Fax +81-78-391-5791
URL http://www.arco.chuo.kobe.jp/
E-Mail office@arco.chuo.kobe.jp

# IN THE UNITED STATES PATENT OFFICE

**U.S. Patent Application No. 09/763,712**
**(U.S. Phase of PCT/JP99/04552)**
**of Nobutaka WAKAMIYA**

I, **Seung-Lim SUNG**, of ARCO PATENT OFFICE at 3rd Fl., Bo-eki Building, 123 Higashi-machi, Chuo-ku, Kobe 650-0031 JAPAN, declare that I am familiar with the Japanese and the English language and, to the best of my knowledge and belief, the attached is a full, true, faithful my prepared English translation of **Japanese Patent Application No. 10-237611 filed on August 24, 1998** which is the priority case in **U.S. Patent Application No. 09/763,712.**

Signature: _____
**Seung-Lim SUNG**

Date: **June 11, 2004**

B11

[DOCUMENT TITLE] Patent Application
[REFERENCE NUMBER] 980824P-03
[FILING DATE] August 24, 1998
[TO] Mr. Takeshi ISAYAMA, Commissioner
[IPC] C12N 15/12
C07K 14/47
C07K 14/435
[TITLE OF INVENTION] Novel Collectin
[NUMBER OF CLAIMS] 19
[INVENTOR]
[ADDRESS] 9-20, Oike 1 chome, Ibaraki-shi, Osaka
[NAME] Nobutaka WAKAMIYA
[APPLICANT]
[ID CODE] 000238201
[NAME] FUSO PHARMACEUTICAL INDUSTRIES, LTD.
[AGENT]
[ID CODE] 100065868
[PATENT ATTORNEY]
[NAME] Yoshihiro SUMIDA
[PHONE NUMBER] 078-321-8822
[APPOINTED AGENT]
[ID CODE] 100088960
[PATENT ATTORNEY]
[NAME] Satoru TAKAISHI
[PHONE NUMBER] 078-321-8822
[APPOINTED AGENT]
[ID CODE] 100106242
[PATENT ATTORNEY]
[NAME] Yasuyuki FURUKAWA
[PHONE NUMBER] 078-321-8822
[APPOINTED AGENT]
[ID CODE] 100107940
[PATENT ATTORNEY]
[NAME] Kengo OKA
[PHONE NUMBER] 078-321-8822

[OFFICIAL FEE]

[ACCOUT NUMBER] 006220

[PAID FEE] 21000 yen

[LIST OF THE FILED DOCUMENTS]

[DOCUMENT TITLE] SPECIFICATION 1

[DOCUMENT TITLE] DRAWINGS 1

[DOCUMENT TITLE] ABSTRACT 1

[ID CODE ON GENERAL POWER OF ATTORNEY] 9705390

[PROOF] Necessary

[DOCUMENT TITLE] SPECIFICATION

[TITLE OF INVENTION] Novel Collectin

[CLAIMS]

[CLAIM 1] A polynucleotide comprising the base sequence that encodes a protein consisting of the amino acid sequence ($206^{th}$-$547^{th}$ residues in SEQ ID NO:2) of:

**Met-Gln-Gln-Asp-Leu-Met-Arg-Ser-Arg-Leu-Asp-Thr-Glu-Val-Ala-Asn-**

**Leu-Ser-Val-Ile-Met-Glu-Glu-Met-Lys-Leu-Val-Asp-Ser-Lys-His-Gly-**

**Gln-Leu-Ile-Lys-Asn-Phe-Thr-Ile-Leu-Gln-Gly-Pro-Pro-Gly-Pro-Arg-**

**Gly-Pro-Arg-Gly-Asp-Arg-Gly-Ser-Gln-Gly-Pro-Pro-Gly-Pro-Thr-Gly-**

**Asn-Lys-Gly-Gln-Lys-Gly-Glu-Lys-Gly-Glu-Pro-Gly-Pro-Pro-Gly-Pro-**

**Ala-Gly-Glu-Arg-Gly-Pro-Ile-Gly-Pro-Ala-Gly-Pro-Pro-Gly-Glu-Arg-**

**Gly-Gly-Lys-Gly-Ser-Lys-Gly-Ser-Gln-Gly-Pro-Lys-Gly-Ser-Arg-Gly-**

**Ser-Pro-Gly-Lys-Pro-Gly-Pro-Gln-Gly-Pro-Ser-Gly-Asp-Pro-Gly-Pro-**

**Pro-Gly-Pro-Pro-Gly-Lys-Glu-Gly-Leu-Pro-Gly-Pro-Gln-Gly-Pro-Pro-**

**Gly-Phe-Gln-Gly-Leu-Gln-Gly-Thr-Val-Gly-Glu-Pro-Gly-Val-Pro-Gly-**

**Pro-Arg-Gly-Leu-Pro-Gly-Leu-Pro-Gly-Val-Pro-Gly-Met-Pro-Gly-Pro-**

**Lys-Gly-Pro-Pro-Gly-Pro-Pro-Gly-Pro-Ser-Gly-Ala-Val-Val-Pro-Leu-**

**Ala-Leu-Gln-Asn-Glu-Pro-Thr-Pro-Ala-Pro-Glu-Asp-Asn-Gly-Cys-Pro-**

**Pro-His-Trp-Lys-Asn-Phe-Thr-Asp-Lys-Cys-Tyr-Tyr-Phe-Ser-Val-Glu-**

**Lys-Glu-Ile-Phe-Glu-Asp-Ala-Lys-Leu-Phe-Cys-Glu-Asp-Lys-Ser-Ser-**

**His-Leu-Val-Phe-Ile-Asn-Thr-Arg-Glu-Glu-Gln-Gln-Trp-Ile-Lys-Lys-**

**Gln-Met-Val-Gly-Arg-Glu-Ser-His-Trp-Ile-Gly-Leu-Thr-Asp-Ser-Glu-**

**Arg-Glu-Asn-Glu-Trp-Lys-Trp-Leu-Asp-Gly-Thr-Ser-Pro-Asp-Tyr-Lys-**

**Asn-Trp-Lys-Ala-Gly-Gln-Pro-Asp-Asn-Trp-Gly-His-Gly-His-Gly-Pro-**

**Gly-Glu-Asp-Cys-Ala-Gly-Leu-Ile-Tyr-Ala-Gly-Gln-Trp-Asn-Asp-Phe-**

**Gln-Cys-Glu-Asp-Val-Asn-Asn-Phe-Ile-Cys-Glu-Lys-Asp-Arg-Glu-Thr-**

**Val-Leu-Ser-Ser-Ala-Leu.**

[CLAIM 2] A polynucleotide comprising the base sequence that encodes a protein consisting of the amino acid sequence (229th-547th residues in SEQ ID NO:2) of:

**Met-Lys-Leu-Val-Asp-Ser-Lys-His-Gly-Gln-Leu-Ile-Lys-Asn-Phe-Thr-Ile-Leu-Gln-Gly-Pro-Pro-Gly-Pro-Arg-Gly-Pro-Arg-Gly-Asp-Arg-Gly-Ser-Gln-Gly-Pro-Pro-Gly-Pro-Thr-Gly-Asn-Lys-Gly-Gln-Lys-Gly-Glu-Lys-Gly-Glu-Pro-Gly-Pro-Pro-Gly-Pro-Ala-Gly-Glu-Arg-Gly-Pro-Ile-Gly-Pro-Ala-Gly-Pro-Pro-Gly-Glu-Arg-Gly-Gly-Lys-Gly-Ser-Lys-Gly-Ser-Gln-Gly-Pro-Lys-Gly-Ser-Arg-Gly-Ser-Pro-Gly-Lys-Pro-Gly-Pro-Gln-Gly-Pro-Ser-Gly-Asp-Pro-Gly-Pro-Pro-Gly-Pro-Pro-Gly-Lys-Glu-Gly-Leu-Pro-Gly-Pro-Gln-Gly-Pro-Pro-Gly-Phe-Gln-Gly-Leu-Gln-Gly-Thr-Val-Gly-Glu-Pro-Gly-Val-Pro-Gly-Pro-Arg-Gly-Leu-Pro-Gly-Leu-Pro-Gly-Val-Pro-Gly-Met-Pro-Gly-Pro-Lys-Gly-Pro-Pro-Gly-Pro-Pro-Gly-Pro-Ser-Gly-Ala-Val-Val-Pro-Leu-Ala-Leu-Gln-Asn-Glu-Pro-Thr-Pro-Ala-Pro-Glu-Asp-Asn-Gly-Cys-Pro-Pro-His-Trp-Lys-Asn-Phe-Thr-Asp-Lys-Cys-Tyr-Tyr-Phe-Ser-Val-Glu-Lys-Glu-Ile-Phe-Glu-Asp-Ala-Lys-Leu-Phe-Cys-Glu-Asp-Lys-Ser-Ser-His-Leu-Val-Phe-Ile-Asn-Thr-Arg-Glu-Glu-Gln-Gln-Trp-Ile-Lys-Lys-Gln-Met-Val-Gly-Arg-Glu-Ser-His-Trp-Ile-Gly-Leu-Thr-Asp-Ser-Glu-Arg-Glu-Asn-Glu-Trp-Lys-Trp-Leu-Asp-Gly-Thr-Ser-Pro-Asp-Tyr-Lys-Asn-Trp-Lys-Ala-Gly-Gln-Pro-Asp-Asn-Trp-Gly-His-Gly-His-Gly-Pro-Gly-Glu-Asp-Cys-Ala-Gly-Leu-Ile-Tyr-Ala-Gly-Gln-Trp-Asn-Asp-Phe-Gln-Cys-Glu-Asp-Val-Asn-Asn-Phe-Ile-Cys-Glu-Lys-Asp-Arg-Glu-Thr-Val-Leu-Ser-Ser-Ala-Leu.**

~~[CLAIM 3] The polynucleotide according to Claim 2~~ further comprises in upstream of the first methionine residue (229th residue in SEQ ID NO:2) the amino acid sequence of:

**Met-Glu-Glu** (226th-228th residues in SEQ ID NO:2); or

**Met-Arg-Ser-Arg-Leu-Asp-Thr-Glu-Val-Ala-Asn-Leu-Ser-Val-Ile-Met-Glu-Glu** (211th-228th residues in SEQ ID NO:2).

[CLAIM 4] The polynucleotide according to Claim 2, wherein said protein further comprises in upstream of the first methionine residue (229$^{th}$ residue in SEQ ID NO:2) the amino acid sequence of:

(102$^{nd}$-228$^{th}$ residues in SEQ ID NO:2)

**Met-Glu-Asn-Ile-Thr-Thr-Ile-Ser-Gln-Ala-Asn-Glu-Gln-Asn-Leu-Lys-Asp-Leu-Gln-Asp-Leu-His-Lys-Asp-Ala-Glu-Asn-Arg-Thr-Ala-Ile-Lys-Phe-Asn-Gln-Leu-Glu-Glu-Arg-Phe-Gln-Leu-Phe-Glu-Thr-Asp-Ile-Val-Asn-Ile-Ile-Ser-Asn-Ile-Ser-Tyr-Thr-Ala-His-His-Leu-Arg-Thr-Leu-Thr-Ser-Asn-Leu-Asn-Glu-Val-Arg-Thr-Thr-Cys-Thr-Asp-Thr-Leu-Thr-Lys-His-Thr-Asp-Asp-Leu-Thr-Ser-Leu-Asn-Asn-Thr-Leu-Ala-Asn-Ile-Arg-Leu-Asp-Ser-Val-Ser-Leu-Arg-Met-Gln-Gln-Asp-Leu-Met-Arg-Ser-Arg-Leu-Asp-Thr-Glu-Val-Ala-Asn-Leu-Ser-Val-Ile-Met-Glu-Glu;**

(91$^{st}$-228$^{th}$ residues in SEQ ID NO:2)

**Met-Asn-Ser-Gln-Leu-Asn-Ser-Phe-Thr-Gly-Gln-Met-Glu-Asn-Ile-Thr-Thr-Ile-Ser-Gln-Ala-Asn-Glu-Gln-Asn-Leu-Lys-Asp-Leu-Gln-Asp-Leu-His-Lys-Asp-Ala-Glu-Asn-Arg-Thr-Ala-Ile-Lys-Phe-Asn-Gln-Leu-Glu-Glu-Arg-Phe-Gln-Leu-Phe-Glu-Thr-Asp-Ile-Val-Asn-Ile-Ile-Ser-Asn-Ile-Ser-Tyr-Thr-Ala-His-His-Leu-Arg-Thr-Leu-Thr-Ser-Asn-Leu-Asn-Glu-Val-Arg-Thr-Thr-Cys-Thr-Asp-Thr-Leu-Thr-Lys-His-Thr-Asp-Asp-Leu-Thr-Ser-Leu-Asn-Asn-Thr-Leu-Ala-Asn-Ile-Arg-Leu-Asp-Ser-Val-Ser-Leu-Arg-Met-Gln-Gln-Asp-Leu-Met-Arg-Ser-Arg-Leu-Asp-Thr-Glu-Val-Ala-Asn-Leu-Ser-Val-Ile-Met-Glu-Glu;**

(9$^{th}$-228$^{th}$ residues in SEQ ID NO:2)

**Met-Asn-Leu-Asn-Asn-Leu-Asn-Leu-Thr-Gln-Val-Gln-Gln-Arg-Asn-Leu-Ile-Thr-Asn-Leu-Gln-Arg-Ser-Val-Asp-Asp-Thr-Ser-Gln-Ala-Ile-Gln-Arg-Ile-Lys-Asn-Asp-Phe-Gln-Asn-Leu-Gln-Gln-Val-Phe-Leu-Gln-Ala-**

Lys-Lys-Asp-Thr-Asp-Trp-Leu-Lys-Glu-Lys-Val-Gln-Ser-Leu-Gln-Thr-Leu-Ala-Ala-Asn-Asn-Ser-Ala-Leu-Ala-Lys-Ala-Asn-Asn-Asp-Thr-Leu-Glu-Asp-Met-Asn-Ser-Gln-Leu-Asn-Ser-Phe-Thr-Gly-Gln-Met-Glu-Asn-Ile-Thr-Thr-Ile-Ser-Gln-Ala-Asn-Glu-Gln-Asn-Leu-Lys-Asp-Leu-Gln-Asp-Leu-His-Lys-Asp-Ala-Glu-Asn-Arg-Thr-Ala-Ile-Lys-Phe-Asn-Gln-Leu-Glu-Glu-Arg-Phe-Gln-Leu-Phe-Glu-Thr-Asp-Ile-Val-Asn-Ile-Ile-Ser-Asn-Ile-Ser-Tyr-Thr-Ala-His-His-Leu-Arg-Thr-Leu-Thr-Ser-Asn-Leu-Asn-Glu-Val-Arg-Thr-Thr-Cys-Thr-Asp-Thr-Leu-Thr-Lys-His-Thr-Asp-Asp-Leu-Thr-Ser-Leu-Asn-Asn-Thr-Leu-Ala-Asn-Ile-Arg-Leu-Asp-Ser-Val-Ser-Leu-Arg-Met-Gln-Gln-Asp-Leu-Met-Arg-Ser-Arg-Leu-Asp-Thr-Glu-Val-Ala-Asn-Leu-Ser-Val-Ile-Met-Glu-Glu; or

($1^{st}$-$228^{th}$ residues in SEQ ID NO:2)

Met-Tyr-Ser-His-Asn-Val-Val-Ile-Met-Asn-Leu-Asn-Asn-Leu-Asn-Leu-Thr-Gln-Val-Gln-Gln-Arg-Asn-Leu-Ile-Thr-Asn-Leu-Gln-Arg-Ser-Val-Asp-Asp-Thr-Ser-Gln-Ala-Ile-Gln-Arg-Ile-Lys-Asn-Asp-Phe-Gln-Asn-Leu-Gln-Gln-Val-Phe-Leu-Gln-Ala-Lys-Lys-Asp-Thr-Asp-Trp-Leu-Lys-Glu-Lys-Val-Gln-Ser-Leu-Gln-Thr-Leu-Ala-Ala-Asn-Asn-Ser-Ala-Leu-Ala-Lys-Ala-Asn-Asn-Asp-Thr-Leu-Glu-Asp-Met-Asn-Ser-Gln-Leu-Asn-Ser-Phe-Thr-Gly-Gln-Met-Glu-Asn-Ile-Thr-Thr-Ile-Ser-Gln-Ala-Asn-Glu-Gln-Asn-Leu-Lys-Asp-Leu-Gln-Asp-Leu-His-Lys-Asp-Ala-Glu-Asn-Arg-Thr-Ala-Ile-Lys-Phe-Asn-Gln-Leu-Glu-Glu-Arg-Phe-Gln-Leu-Phe-Glu-Thr-Asp-Ile-Val-Asn-Ile-Ile-Ser-Asn-Ile-Ser-Tyr-Thr-Ala-His-His-Leu-Arg-Thr-Leu-Thr-Ser-Asn-Leu-Asn-Glu-Val-Arg-Thr-Thr-Cys-Thr-Asp-Thr-Leu-Thr-Lys-His-Thr-Asp-Asp-Leu-Thr-Ser-Leu-Asn-Asn-Thr-Leu-Ala-Asn-Ile-Arg-Leu-Asp-Ser-Val-Ser-Leu-Arg-Met-Gln-Gln-Asp-Leu-Met-Arg-Ser-Arg-Leu-Asp-Thr-Glu-Val-Ala-Asn-Leu-Ser-Val-

Ile-Met-Glu-Glu.

[CLAIM 5] A polynucleotide comprising the base sequence (670th-1695th bases in SEQ ID NO:1) of:

atgcaacaag atttgatgag gtcgaggtta gacactgaag tagccaactt atcagtgatt atggaagaaa tgaagctagt agactccaag catggtcagc tcatcaagaa ttttacaata ctacaaggtc caccgggccc caggggtcca agaggtgaca gaggatccca gggacccccc ggcccaactg gcaacaaggg acagaaagga gagaaggggg agcctggacc acctggccct gcgggtgaga gaggcccaat tggaccagct ggtccccccg gagagcgtgg cggcaaagga tctaaaggct cccagggccc caaaggctcc cgtggttccc ctgggaagcc cggccctcag ggccccagtg gggacccagg ccccccgggc ccaccaggca aagagggact ccccggccct cagggccctc ctggcttcca gggacttcag ggcaccgttg gggagcctgg ggtgcctgga cctcggggac tgccaggctt gcctggggta ccaggcatgc caggccccaa gggccccccc ggccctcctg gcccatcagg agcggtggtg cccctggccc tgcagaatga gccaaccccg gcaccggagg acaatggctg cccgcctcac tggaagaact tcacagacaa atgctactat ttttcagttg agaaagaaat ttttgaggat gcaaagcttt tctgtgaaga caagtcttca catcttgttt tcataaacac tagagaggaa cagcaatgga taaaaaaaca gatggtaggg agagagagcc actggatcgg cctcacagac tcagagcgtg aaaatgaatg gaagtggctg gatgggacat ctccagacta caaaaattgg aaagctggac agccggataa ctggggtcat ggccatgggc caggagaaga ctgtgctggg ttgatttatg ctgggcagtg gaacgatttc caatgtgaag acgtcaataa cttcatttgc gaaaaagaca gggagacagt actgtcatct gcatta.

[CLAIM 6] A polynucleotide comprising the base sequence (739th-1695th bases in SEQ ID NO:1) of:

atgaagctag tagactccaa gcatggtcag ctcatcaaga attttacaat actacaaggt ccaccgggcc ccaggggtcc aagaggtgac agaggatccc agggaccccc tggcccaact ggcaacaagg gacagaaagg agagaagggg gagcctggac cacctggccc tgcgggtgag agaggcccaa ttggaccagc tggtcccccc ggagagcgtg gcggcaaagg atctaaaggc

tcccagggcc ccaaaggctc ccgtggttcc cctgggaagc ccggccctca gggccccagt

ggggacccag gccccccggg cccaccaggc aaagagggac tccccggccc tcagggccct

cctggcttcc agggacttca gggcaccgtt ggggagcctg gggtgcctgg acctcgggga

ctgccaggct tgcctggggt accaggcatg ccaggcccca agggcccccc cggccctcct

ggcccatcag gagcggtggt gcccctggcc ctgcagaatg agccaacccc ggcaccggag

gacaatggct gcccgcctca ctggaagaac ttcacagaca aatgctacta tttttcagtt

gagaaagaaa tttttgagga tgcaaagctt ttctgtgaag acaagtcttc acatcttgtt

ttcataaaca ctagagagga acagcaatgg ataaaaaaac agatggtagg gagagagagc

cactggatcg gcctcacaga ctcagagcgt gaaaatgaat ggaagtggct ggatgggaca

tctccagact acaaaaattg gaaagctgga cagccggata actggggtca tggccatggg

ccaggagaag actgtgctgg gttgatttat gctgggcagt ggaacgattt ccaatgtgaa

gacgtcaata acttcatttg cgaaaaagac agggagacag tactgtcatc tgcatta.

【CLAIM 7】 The polynucleotide according to Claim 6 further comprises, in 5` upstream of said base sequence, the base sequence of:

atggaagaa (730$^{th}$-738$^{th}$ bases in SEQ ID NO:1); or

atgaggtcga ggttagacac tgaagtagcc aacttatcag tgattatgga agaa

(685$^{th}$-738$^{th}$ bases in SEQ ID NO:1).

【CLAIM 8】 The polynucleotide according to Claim 6 further comprises, in 5` upstream of said base sequence, the base sequence of:

(358$^{th}$-738$^{th}$ bases in SEQ ID NO:1)

atggagaaca tcaccactat ctctcaagcc aacgagcaga acctgaaaga cctgcaggac

ttacacaaag atgcagagaa tagaacagcc atcaagttca accaactgga ggaacgcttc

cagctctttg agacggatat tgtgaacatc attagcaata tcagttacac agcccaccac

ctgcggacgc tgaccagcaa tctaaatgaa gtcaggacca cttgcacaga tacccttacc

aaacacacag atgatctgac ctccttgaat aataccctgg ccaacatccg tttggattct

gtttctctca ggatgcaaca agatttgatg aggtcgaggt tagacactga agtagccaac

ttatcagtga ttatggaaga a;

(325th-738th bases in SEQ ID NO:1)

atgaacagcc agctcaactc attcacaggt cagatggaga acatcaccac tatctctcaa

gccaacgagc agaacctgaa agacctgcag gacttacaca aagatgcaga gaatagaaca

gccatcaagt tcaaccaact ggaggaacgc ttccagctct ttgagacgga tattgtgaac

atcattagca atatcagtta cacagcccac cacctgcgga cgctgaccag caatctaaat

gaagtcagga ccacttgcac agataccctt accaaacaca cagatgatct gacctccttg

aataataccc tggccaacat ccgtttggat tctgtttctc tcaggatgca acaagatttg

atgaggtcga ggttagacac tgaagtagcc aacttatcag tgattatgga agaa;

(79th-738th bases in SEQ ID NO:1)

atgaacctca acaacctgaa cctgacccag gtgcagcaga ggaacctcat cacgaatctg

cagcggtctg tggatgacac aagccaggct atccagcgaa tcaagaacga ctttcaaaat

ctgcagcagg tttttcttca agccaagaag gacacggatt ggctgaagga gaaagtgcag

agcttgcaga cgctggctgc caacaactct gcgttggcca aagccaacaa cgacaccctg

gaggatatga acagccagct caactcattc acaggtcaga tggagaacat caccactatc

tctcaagcca acgagcagaa cctgaaagac ctgcaggact tacacaaaga tgcagagaat

agaacagcca tcaagttcaa ccaactggag gaacgcttcc agctctttga gacggatatt

gtgaacatca ttagcaatat cagttacaca gcccaccacc tgcggacgct gaccagcaat

ctaaatgaag tcaggaccac ttgcacagat acccttacca aacacacaga tgatctgacc

tccttgaata ataccctggc caacatccgt ttggattctg tttctctcag gatgcaacaa

gatttgatga ggtcgaggtt agacactgaa gtagccaact tatcagtgat tatggaagaa;

(55th-738th bases in SEQ ID NO:1)

atgtattctc ataatgtggt catcatgaac ctcaacaacc tgaacctgac ccaggtgcag

cagaggaacc tcatcacgaa tctgcagcgg tctgtggatg acacaagcca ggctatccag

cgaatcaaga acgactttca aaatctgcag caggtttttc ttcaagccaa gaaggacacg

gattggctga aggagaaagt gcagagcttg cagacgctgg ctgccaacaa ctctgcgttg

gccaaagcca acaacgacac cctggaggat atgaacagcc agctcaactc attcacaggt cagatggaga acatcaccac tatctctcaa gccaacgagc agaacctgaa agacctgcag gacttacaca aagatgcaga gaatagaaca gccatcaagt tcaaccaact ggaggaacgc ttccagctct ttgagacgga tattgtgaac atcattagca atatcagtta cacagcccac cacctgcgga cgctgaccag caatctaaat gaagtcagga ccacttgcac agataccctt accaaacaca cagatgatct gacctccttg aataataccc tggccaacat ccgtttggat tctgtttctc tcaggatgca acaagatttg atgaggtcga ggttagacac tgaagtagcc aacttatcag tgattatgga agaa; or

($1^{st}$-$738^{th}$ bases in SEQ ID NO:1)

gtcacgaatc tgcagcaaga taccagcgtg ctccagggca atctgcagaa ccaaatgtat tctcataatg tggtcatcat gaacctcaac aacctgaacc tgacccaggt gcagcagagg aacctcatca cgaatctgca gcggtctgtg gatgacacaa gccaggctat ccagcgaatc aagaacgact ttcaaaatct gcagcaggtt tttcttcaag ccaagaagga cacggattgg ctgaaggaga aagtgcagag cttgcagacg ctggctgcca acaactctgc gttggccaaa gccaacaacg acaccctgga ggatatgaac agccagctca actcattcac aggtcagatg gagaacatca ccactatctc tcaagccaac gagcagaacc tgaaagacct gcaggactta cacaaagatg cagagaatag aacagccatc aagttcaacc aactggagga acgcttccag ctctttgaga cggatattgt gaacatcatt agcaatatca gttacacagc ccaccacctg cggacgctga ccagcaatct aaatgaagtc aggaccactt gcacagatac ccttaccaaa cacacagatg atctgacctc cttgaataat accctggcca acatccgttt ggattctgtt tctctcagga tgcaacaaga tttgatgagg tcgaggttag acactgaagt agccaactta tcagtgatta tggaagaa.

[CLAIM 9] The polynucleotide according to any of Claims 5-8 further comprises, in 3` downstream of said base sequence, the base sequence of:

(1696$^{th}$-2024$^{th}$ bases in SEQ ID NO:1)

taacggactg tgatgggatc acatgagcaa attttcagct ctcaaaggca aaggacactc ctttctaatt gcatcacctt ctcatcagat tgaaaaaaaa aaaagcactg aaaaccaatt actgaaaaaa aattgacagc tagtgttttt taccatccgt cattacccaa agacttggga actaaaatgt tccccagggt gatatgctga ttttcattgt gcacatggac tgaatcacat agattctcct ccgtcagtaa ccgtgcgatt atacaaatta tgtcttccaa agtatggaac actccaatca gaaaaaggtt atcatcccg.

[CLAIM 10] A polynucleotide which encode a novel collectin and can hybridize under a stringent condition with a probe of an amplification product from PCR reaction performed using primers having the base sequences of:

caatctgatgagaaggtgatg (SEQ ID NO: 4) and

acgaggggctggatgggacat (SEQ ID NO: 5).

[CLAIM 11] A polynucleotide which can hybridize under a stringent condition with the polynucleotide according to any of Claims 1-10, wherein a protein encoded by the polynucleotide is a novel collectin comprising: (1) $Ca^{2+}$-dependent carbohydrate recognition domain (CRD), and (2) collagen-like region.

[CLAIM 12] The polynucleotide according to any of Claims 1-11, wherein said polynucleotide is cDNA.

[CLAIM 13] Collectin protein comprising the amino acid sequence encoded by the polynucleotide according to any of Claims 5-12.

[CLAIM 14] Collectin protein comprising the amino acid sequence of:

(206$^{th}$ -547$^{th}$ residues in SEQ ID NO: 2)

Met-Gln-Gln-Asp-Leu-Met-Arg-Ser-Arg-Leu-Asp-Thr-Glu-Val-Ala-Asn-Leu-Ser-Val-Ile-Met-Glu-Glu-Met-Lys-Leu-Val-Asp-Ser-Lys-His-Gly-Gln-Leu-Ile-Lys-Asn-Phe-Thr-Ile-Leu-Gln-Gly-Pro-Pro-Gly-Pro-Arg-Gly-Pro-Arg-Gly-Asp-Arg-Gly-Ser-Gln-Gly-Pro-Pro-Gly-Pro-Thr-Gly-Asn-Lys-Gly-Gln-Lys-Gly-Glu-Lys-Gly-Glu-Pro-Gly-Pro-Pro-Gly-Pro-

Ala-Gly-Glu-Arg-Gly-Pro-Ile-Gly-Pro-Ala-Gly-Pro-Pro-Gly-Glu-Arg-Gly-Gly-Lys-Gly-Ser-Lys-Gly-Ser-Gln-Gly-Pro-Lys-Gly-Ser-Arg-Gly-Ser-Pro-Gly-Lys-Pro-Gly-Pro-Gln-Gly-Pro-Ser-Gly-Asp-Pro-Gly-Pro-Pro-Gly-Pro-Pro-Gly-Lys-Glu-Gly-Leu-Pro-Gly-Pro-Gln-Gly-Pro-Pro-Gly-Phe-Gln-Gly-Leu-Gln-Gly-Thr-Val-Gly-Glu-Pro-Gly-Val-Pro-Gly-Pro-Arg-Gly-Leu-Pro-Gly-Leu-Pro-Gly-Val-Pro-Gly-Met-Pro-Gly-Pro-Lys-Gly-Pro-Pro-Gly-Pro-Pro-Gly-Pro-Ser-Gly-Ala-Val-Val-Pro-Leu-Ala-Leu-Gln-Asn-Glu-Pro-Thr-Pro-Ala-Pro-Glu-Asp-Asn-Gly-Cys-Pro-Pro-His-Trp-Lys-Asn-Phe-Thr-Asp-Lys-Cys-Tyr-Tyr-Phe-Ser-Val-Glu-Lys-Glu-Ile-Phe-Glu-Asp-Ala-Lys-Leu-Phe-Cys-Glu-Asp-Lys-Ser-Ser-His-Leu-Val-Phe-Ile-Asn-Thr-Arg-Glu-Glu-Gln-Gln-Trp-Ile-Lys-Lys-Gln-Met-Val-Gly-Arg-Glu-Ser-His-Trp-Ile-Gly-Leu-Thr-Asp-Ser-Glu-Arg-Glu-Asn-Glu-Trp-Lys-Trp-Leu-Asp-Gly-Thr-Ser-Pro-Asp-Tyr-Lys-Asn-Trp-Lys-Ala-Gly-Gln-Pro-Asp-Asn-Trp-Gly-His-Gly-His-Gly-Pro-Gly-Glu-Asp-Cys-Ala-Gly-Leu-Ile-Tyr-Ala-Gly-Gln-Trp-Asn-Asp-Phe-Gln-Cys-Glu-Asp-Val-Asn-Asn-Phe-Ile-Cys-Glu-Lys-Asp-Arg-Glu-Thr-Val-Leu-Ser-Ser-Ala-Leu.

【CLAIM 15】 Collectin protein comprising the amino acid sequence of:

(229$^{th}$-547$^{th}$ residues in SEQ ID NO: 2)

Met-Lys-Leu-Val-Asp-Ser-Lys-His-Gly-Gln-Leu-Ile-Lys-Asn-Phe-Thr-Ile-Leu-Gln-Gly-Pro-Pro-Gly-Pro-Arg-Gly-Pro-Arg-Gly-Asp-Arg-Gly-Ser-Gln-Gly-Pro-Pro-Gly-Pro-Thr-Gly-Asn-Lys-Gly-Gln-Lys-Gly-Glu-Lys-Gly-Glu-Pro-Gly-Pro-Pro-Gly-Pro-Ala-Gly-Glu-Arg-Gly-Pro-Ile-Gly-Pro-Ala-Gly-Pro-Pro-Gly-Glu-Arg-Gly-Gly-Lys-Gly-Ser-Lys-Gly-Ser-Gln-Gly-Pro-Lys-Gly-Ser-Arg-Gly-Ser-Pro-Gly-Lys-Pro-Gly-Pro-Gln-Gly-Pro-Ser-Gly-Asp-Pro-Gly-Pro-Pro-Gly-Pro-Pro-Gly-Lys-Glu-

Gly-Leu-Pro-Gly-Pro-Gln-Gly-Pro-Pro-Gly-Phe-Gln-Gly-Leu-Gln-Gly-Thr-Val-Gly-Glu-Pro-Gly-Val-Pro-Gly-Pro-Arg-Gly-Leu-Pro-Gly-Leu-Pro-Gly-Val-Pro-Gly-Met-Pro-Gly-Pro-Lys-Gly-Pro-Pro-Gly-Pro-Pro-Gly-Pro-Ser-Gly-Ala-Val-Val-Pro-Leu-Ala-Leu-Gln-Asn-Glu-Pro-Thr-Pro-Ala-Pro-Glu-Asp-Asn-Gly-Cys-Pro-Pro-His-Trp-Lys-Asn-Phe-Thr-Asp-Lys-Cys-Tyr-Tyr-Phe-Ser-Val-Glu-Lys-Glu-Ile-Phe-Glu-Asp-Ala-Lys-Leu-Phe-Cys-Glu-Asp-Lys-Ser-Ser-His-Leu-Val-Phe-Ile-Asn-Thr-Arg-Glu-Glu-Gln-Gln-Trp-Ile-Lys-Lys-Gln-Met-Val-Gly-Arg-Glu-Ser-His-Trp-Ile-Gly-Leu-Thr-Asp-Ser-Glu-Arg-Glu-Asn-Glu-Trp-Lys-Trp-Leu-Asp-Gly-Thr-Ser-Pro-Asp-Tyr-Lys-Asn-Trp-Lys-Ala-Gly-Gln-Pro-Asp-Asn-Trp-Gly-His-Gly-His-Gly-Pro-Gly-Glu-Asp-Cys-Ala-Gly-Leu-Ile-Tyr-Ala-Gly-Gln-Trp-Asn-Asp-Phe-Gln-Cys-Glu-Asp-Val-Asn-Asn-Phe-Ile-Cys-Glu-Lys-Asp-Arg-Glu-Thr-Val-Leu-Ser-Ser-Ala-Leu.

【CLAIM 16】 The collectin protein according to Claim 15 further comprises, in upstream of of the first methionine residue (229$^{th}$ residue in SEQ ID NO: 2), the amino acid sequence of:

Met-Glu-Glu (226$^{th}$-228$^{th}$ residues in SEQ ID NO: 2); or

Met-Arg-Ser-Arg-Leu-Asp-Thr-Glu-Val-Ala-Asn-Leu-Ser-Val-Ile-Met-Glu-Glu (211$^{th}$-228$^{th}$ residues in SEQ ID NO: 2).

【CLAIM 17】 The collectin protein according to Claim 15 further comprises, in upstream of of the first methionine residue (229$^{th}$ residue in SEQ ID NO: 2), the amino acid sequence of: (102$^{nd}$-228$^{th}$ residues in SEQ ID NO: 2)

Met-Glu-Asn-Ile-Thr-Thr-Ile-Ser-Gln-Ala-Asn-Glu-Gln-Asn-Leu-Lys-Asp-Leu-Gln-Asp-Leu-His-Lys-Asp-Ala-Glu-Asn-Arg-Thr-Ala-Ile-Lys-Phe-Asn-Gln-Leu-Glu-Glu-Arg-Phe-Gln-Leu-Phe-Glu-Thr-Asp-Ile-Val-Asn-Ile-Ile-Ser-Asn-Ile-Ser-Tyr-Thr-Ala-His-His-Leu-Arg-Thr-Leu-Thr-Ser-Asn-Leu-Asn-Glu-Val-Arg-Thr-Thr-Cys-Thr-Asp-Thr-Leu-Thr-

Lys-His-Thr-Asp-Asp-Leu-Thr-Ser-Leu-Asn-Asn-Thr-Leu-Ala-Asn-Ile-Arg-Leu-Asp-Ser-Val-Ser-Leu-Arg-Met-Gln-Gln-Asp-Leu-Met-Arg-Ser-Arg-Leu-Asp-Thr-Glu-Val-Ala-Asn-Leu-Ser-Val-Ile-Met-Glu-Glu;

(91st-228th residues in SEQ ID NO: 2)

Met-Asn-Ser-Gln-Leu-Asn-Ser-Phe-Thr-Gly-Gln-Met-Glu-Asn-Ile-Thr-Thr-Ile-Ser-Gln-Ala-Asn-Glu-Gln-Asn-Leu-Lys-Asp-Leu-Gln-Asp-Leu-His-Lys-Asp-Ala-Glu-Asn-Arg-Thr-Ala-Ile-Lys-Phe-Asn-Gln-Leu-Glu-Glu-Arg-Phe-Gln-Leu-Phe-Glu-Thr-Asp-Ile-Val-Asn-Ile-Ile-Ser-Asn-Ile-Ser-Tyr-Thr-Ala-His-His-Leu-Arg-Thr-Leu-Thr-Ser-Asn-Leu-Asn-Glu-Val-Arg-Thr-Thr-Cys-Thr-Asp-Thr-Leu-Thr-Lys-His-Thr-Asp-Asp-Leu-Thr-Ser-Leu-Asn-Asn-Thr-Leu-Ala-Asn-Ile-Arg-Leu-Asp-Ser-Val-Ser-Leu-Arg-Met-Gln-Gln-Asp-Leu-Met-Arg-Ser-Arg-Leu-Asp-Thr-Glu-Val-Ala-Asn-Leu-Ser-Val-Ile-Met-Glu-Glu;

(9th-228th residues in SEQ ID NO: 2)

Met-Asn-Leu-Asn-Asn-Leu-Asn-Leu-Thr-Gln-Val-Gln-Gln-Arg-Asn-Leu-Ile-Thr-Asn-Leu-Gln-Arg-Ser-Val-Asp-Asp-Thr-Ser-Gln-Ala-Ile-Gln-Arg-Ile-Lys-Asn-Asp-Phe-Gln-Asn-Leu-Gln-Gln-Val-Phe-Leu-Gln-Ala-Lys-Lys-Asp-Thr-Asp-Trp-Leu-Lys-Glu-Lys-Val-Gln-Ser-Leu-Gln-Thr-Leu-Ala-Ala-Asn-Asn-Ser-Ala-Leu-Ala-Lys-Ala-Asn-Asn-Asp-Thr-Leu-Glu-Asp-Met-Asn-Ser-Gln-Leu-Asn-Ser-Phe-Thr-Gly-Gln-Met-Glu-Asn-Ile-Thr-Thr-Ile-Ser-Gln-Ala-Asn-Glu-Gln-Asn-Leu-Lys-Asp-Leu-Gln-Asp-Leu-His-Lys-Asp-Ala-Glu-Asn-Arg-Thr-Ala-Ile-Lys-Phe-Asn-Gln-Leu-Glu-Glu-Arg-Phe-Gln-Leu-Phe-Glu-Thr-Asp-Ile-Val-Asn-Ile-Ile-Ser-Asn-Ile-Ser-Tyr-Thr-Ala-His-His-Leu-Arg-Thr-Leu-Thr-Ser-Asn-Leu-Asn-Glu-Val-Arg-Thr-Thr-Cys-Thr-Asp-Thr-Leu-Thr-Lys-His-Thr-Asp-Asp-Leu-Thr-Ser-Leu-Asn-Asn-Thr-Leu-Ala-Asn-Ile-Arg-Leu-Asp-

Ser-Val-Ser-Leu-Arg-Met-Gln-Gln-Asp-Leu-Met-Arg-Ser-Arg-Leu-Asp-Thr-Glu-Val-Ala-Asn-Leu-Ser-Val-Ile-Met-Glu-Glu; or

($1^{st}$-$228^{th}$ residues in SEQ ID NO: 2)
Met-Tyr-Ser-His-Asn-Val-Val-Ile-Met-Asn-Leu-Asn-Asn-Leu-Asn-Leu-Thr-Gln-Val-Gln-Gln-Arg-Asn-Leu-Ile-Thr-Asn-Leu-Gln-Arg-Ser-Val-Asp-Asp-Thr-Ser-Gln-Ala-Ile-Gln-Arg-Ile-Lys-Asn-Asp-Phe-Gln-Asn-Leu-Gln-Gln-Val-Phe-Leu-Gln-Ala-Lys-Lys-Asp-Thr-Asp-Trp-Leu-Lys-Glu-Lys-Val-Gln-Ser-Leu-Gln-Thr-Leu-Ala-Ala-Asn-Asn-Ser-Ala-Leu-Ala-Lys-Ala-Asn-Asn-Asp-Thr-Leu-Glu-Asp-Met-Asn-Ser-Gln-Leu-Asn-Ser-Phe-Thr-Gly-Gln-Met-Glu-Asn-Ile-Thr-Thr-Ile-Ser-Gln-Ala-Asn-Glu-Gln-Asn-Leu-Lys-Asp-Leu-Gln-Asp-Leu-His-Lys-Asp-Ala-Glu-Asn-Arg-Thr-Ala-Ile-Lys-Phe-Asn-Gln-Leu-Glu-Glu-Arg-Phe-Gln-Leu-Phe-Glu-Thr-Asp-Ile-Val-Asn-Ile-Ile-Ser-Asn-Ile-Ser-Tyr-Thr-Ala-His-His-Leu-Arg-Thr-Leu-Thr-Ser-Asn-Leu-Asn-Glu-Val-Arg-Thr-Thr-Cys-Thr-Asp-Thr-Leu-Thr-Lys-His-Thr-Asp-Asp-Leu-Thr-Ser-Leu-Asn-Asn-Thr-Leu-Ala-Asn-Ile-Arg-Leu-Asp-Ser-Val-Ser-Leu-Arg-Met-Gln-Gln-Asp-Leu-Met-Arg-Ser-Arg-Leu-Asp-Thr-Glu-Val-Ala-Asn-Leu-Ser-Val-Ile-Met-Glu-Glu.

[CLAIM 18] The collectin protein according to any of Claims 13-17 which is from human.

[CLAIM 19] The collectin protein which consists of the amino acid sequence comprising deletion, substitution and/or addition of one or more amino acid/s in the collectin protein according to any of Claims 13-18, and comprises (1) $Ca^{2+}$-dependent carbohydrate recognition domain (CRD), and (2) collagen-like region.

[DETAILED DESCRIPTION OF THE INVENTION]

[0001]

[TECHNICAL FIELD WHERE THE INVENTION BELONGS]

The present invention relates to a novel collectin which is useful for investigating mechanisms of biological defense system, and is expected to be applied for utilizing as materials for medicines because it may have physiological activities including anti-viral activities and the like.

【0002】

[PRIOR ART]

Collectin is a generic name of proteins having $Ca^{2+}$-dependent carbohydrate recognition region (CRD) and collagen-like region, and the member of these proteins is conceived to involve in basic immunity systems against a wide spectrum of microorganisms such as bacteria and viruses.

【0003】

The collectins that have been identified heretofore include mannan-binding protein (MBP), surfactant protein A (SP-A), surfactant protein D (SP-D), conglutinin and the like. These collectins are known to be constituted from basic structures, as illustrated in Fig. 1(a), comprising unique regions of: (1) $Ca^{2+}$-dependent carbohydrate recognition domain (CRD), and (2) collagen-like region [Malhortra *et al.*, *Eur.J.Immunol.* Vol.22, 1437-1445, 1992], and a subunit may be formed from the three basic structures through making a triple helix in the collagen-like region, besides, such subunits may form an oligomer, e.g., trimer, tetramer and hexamer.

【0004】

In vertebrates, mechanisms involving cellular immune responses and specific antibody reactions are considered as dominant host-defense systems against inversion of the pathogenic bacteria, viruses and the like. Recently, involvement in nonspecific immune responses of the lectins such as conglutinin has been suggested, for example, it was reported that the lectins may play important roles in neutralizing and removing the various microorganisms in infants having insufficient maternal antibodies and undeveloped specific defense systems [Super *et al.*, *Lancet*, Vol.2, 1236-1239, 1989]. Moreover, with respect to the roles of the lectins in the biological host-defense systems, it was reported that the host becomes liable to be infected by, for example, a reduction of the MBP concentration in blood due to genetic mutation of the MBP gene [Sumiya *et al.*, *Lancet*, Vol.337, 1569-1570, 1991].

【0005】

The present inventors previously found that the conglutinin and the mannan-binding protein can inhibit infection and hemagglutination activity of H1 and H3 Type Influenza A viruses [Wakamiya et al., *Glycoconjugate J.*, Vol.8, 235, 1991; Wakamiya *et al.*, *Biochem. Biophys. Res. Comm.*, Vol.187, 1270-1278, 1992].

【0006】

Thereafter, the present inventors isolated a cDNA clone encoding the conglutinin, and found that closer correlation may exist between the conglutinin gene and various surfactant protein D gene [Suzuki *et al.*, *Biochem. Biophys. Res. Comm.*, Vol.191, 335-342, 1993].

【0007】

Thus, the collectin has been expected to be useful in investigating mechanisms of biological defense, and useful as a physiologically active medical material. Therefore, discovery of novel molecular species belonging to this family would largely contribute to various medical fields including therapy of infectious diseases, as well as biological fields.

【0008】

[PROBLEMS TO BE SOLVED BY THE INVENTION]

The present invention was accomplished in consideration of the aforementioned state of art, and an object of the invention is to provide a novel collectin which can be expected to exhibit physiological activities such as anti-bacterial, anti-viral activities, especially in human body.

【0009】

[ELEMENTS TO SOLVE THE PROBLEMS]

Accordingly, the present invention provides novel collectin gene and protein having characteristic structures of the collectins, which are distinct from those reported in the art, as follows:

[1] A polynucleotide comprising a nucleotide sequence which encodes a protein having an amino acid sequence of $206^{th}$-$547^{th}$ residues in SEQ ID NO: 2;

[2] A polynucleotide comprising a nucleotide sequence which encodes a protein having an amino acid sequence of $229^{th}$-$547^{th}$ residues in SEQ ID NO: 2;

[3] A polynucleotide comprising a base sequence of $670^{th}$-$1695^{th}$ bases in SEQ ID NO: 1;

[4] A polynucleotide comprising a nucleotide sequence of $739^{th}$-$1695^{th}$ bases in SEQ ID NO: 1;

[5] A polynucleotide which encodes a collectin protein and can hybridize under a stringent condition with a probe of an amplification product from PCR reaction performed using primers having base sequences set out in SEQ ID NOs: 4 and 5;

[6] A polynucleotide which can hybridize under a stringent condition with any of the polynucleotides [1]-[5], wherein the protein encoded by the polynucleotide is a human collectin protein comprising (1) $Ca^{2+}$-dependent carbohydrate recognition domain (CRD), and (2) collagen-like region;

[7] A collectin protein encoded by any of polynucleotides [3]-[6];

[8] A collectin protein comprising an amino acid sequence of $206^{th}$-$547^{th}$ residues in SEQ ID NO:2;

[9] A collectin protein comprising an amino acid sequence of $229^{th}$-$547^{th}$ residues in SEQ ID NO:2;

[10] The collectin protein according to any of the collectin proteins [7]-[9] which consists of the amino acid sequence that comprises deletion, substitution and/or addition of one or more amino acid/s, and the amino acid sequence comprises (1) $Ca^{2+}$-dependent carbohydrate recognition domain (CRD), and (2) collagen-like region.

【0010】

【EMBODIMENT】

In the preferred embodiment of the present invention, the polynucleotides [1]-[6] of the present invention may preferably be cDNA.

【0011】

The polynucleotide [2] comprises a nucleotide sequence encoding a protein having at least an amino acid sequence of

229th-547th residues in SEQ ID NO:2, however, additional nucleotide sequence such as those encoding a protein having an amino acid sequence such as 226th-228th residues or 211th-228th residues of SEQ ID NO: 2, or 102nd-228th residues, 91st-228th residues, 9th-228th residues, 1st-228th residues of SEQ ID NO: 2 may be contained upstream of the first methionine residue.

[0012]

Moreover, the polynucleotide [4] may further comprise, in 5' upstream thereof, a base sequence of 730th-738th bases or 685th-738th bases, or 358th-738th bases, 325th-738th bases, 79th-738th bases, 55th-738th bases or 1st-738th bases in SEQ ID NO: 1.

[0013]

Additionally, the polynucleotide [3] or [4] may comprise, in 3' upstream thereof, a base sequence of 1696th-2024th bases in SEQ ID NO:1.

[0014]

Further, taking into account of the usefulness as physiologically active medical materials, the proteins [7]-[10] of the present invention may preferably be from human, because it can be expected to exhibit anti-bacterial, anti-viral activities and the like in human body. Accordingly, the present invention contemplates a collectin protein derived from human, and upon examination of various human tissues, expression of a novel collectin was suggested, which was conceived to be useful.

[0015]

Further, the collectin protein [9] may preferably comprise at least an amino acid sequence of 229th-547th residues in SEQ ID NO: 2, and an amino acid sequence, for example, 226th-228th residues or 211th-228th residues in SEQ ID NO: 2, or 102nd-228th residues, 91st-228th residues, 9th-228th residues, 1st-228th residues in SEQ ID NO: 2 or the like may be additionally included in upstream of the first methionine residue.

[0016]

The stringent hybridization condition referred to in [5]-[6] may include, for example, the serial steps of:

prehybridization in a solution of 5 x SSC (prepared by diluting 20 x SSC (3 M NaCl, 0.3 M sodium citrate)), 1% blocking agent (Boehringer Mannheim), 0.1% N-lauroyl sarcosine, and 0.02% SDS, at 68 C for one hour; hybridization in a solution of 5 x SSC, 1% blocking agent, 0.1% N-lauroyl sarcosine, and 0.02% SDS containing cDNA probes (10 ng/ml), at 55 C for 16 hours; washing twice in a solution of 2 x SSC/0.1% SDS for 5 minutes; and washing twice in a solution of 0.5 x SSC/0.1% SDS at 55 C for 15 minutes, but some modifications/alterations of these conditions such as the concentration of the solution, incubation temperature and time may be made on the basis of the knowledge in the art.

[0017]

Further, deletion, substitution and/or addition of one or more amino acids as referred to in [10] above may be those which does not result in great changes of hydrophilic/hydrophobic, acidic/basic nature, functional groups in the collectin protein, and may not bring much alteration on the properties by (1) $Ca^{2+}$-dependent carbohydrate recognition domain (CRD) and (2) collagen-like region. On the basis of the amino acid sequences and structures of the proteins belonging to the collectin families reported in the art, for example, deletion, substitution and/or addition of 1-10 amino acid residue/s in (1) $Ca^{2+}$-dependent carbohydrate recognition domain (CRD), and 1-100, preferably 1-15 amino acid residue/s in (2) collagen-like region may be allowed.

[0018]

Although the present invention will be described in more detail along with the following examples, as a matter of course, the present invention should not be interpreted narrower based on the disclosure of the following examples.

[0019]

The Examples demonstrate the search on EST database (Example 1); preparation of the probe for screening (Example 2); screening of cDNA library from human placenta (Example 3); sequencing of the base sequence of the novel collectin (Example 4); genomic Southern analysis of the novel collectin (Example

5); Northern analysis of the novel collectin with various human tissues (Example 6); genomic Southern analysis of the novel collectin with tissues from various animal species (Example 7); and genetic analysis of the novel collectin (Example 8).

[0020]

Example 1: Search on EST Database

Highly conserved regions between molecules of the known collectins, i.e., human MBP, human SP-A and human SP-D were searched by comparing the amino acid sequences thereof (see Figures 2 and 3, in which amino acid residues which were recognized to be homologous between those proteins were boxed). As a result thereof, it was suggested that the region consisting of 27 amino acids, namely from 220$^{th}$ amino acid to 246$^{th}$ amino acid in human MBP sequence (shown in Figure 3, reversed characters, SEQ ID NO:6), was highly homologous, some consensus sequences corresponding to this region were therefore prepared, and searches on EST (Expressed Sequence Tags) database were conducted with such sequences. For this search, the EST database including 676750 sequences published on October 11, 1996 was used.

[0021]

Consequently, some data comprising highly homologous amino acid sequences with the sequence of the 27 amino acid described above were obtained. The amino acid sequences in the obtained data were further searched with GenBank/EST database, then they were determined as to whether they were from known or unknown genes, and it was confirmed that, when the the amino acid sequence of **Glu-Lys-Cys-Val-Glu-Met-Tyr-Thr-Asp-Gly-Lys-Trp-Asn-Asp-Arg-Asn-Cys-Leu-Gln-Ser-Arg-Leu-Ala-Ile-Cys-Glu-Phe** (SEQ ID NO: 3) was used as a consensus sequence, there was two data (Registration Nos. W72977 and R74387) having highly homologous but with unknown base sequences. These were respectively from placenta and from fetal heart, and were clones indicating portions of base sequences of the novel collectin.

[0022]

Thereafter, the clone from fetal heart (I.M.A.G.E. Consortium Clone ID 34472) was purchased from ATCC (American Type Culture Collection), and utilized to prepare a screening probe for obtaining a novel collectin.

[0023]

Example 2: Preparation of Probe for Screening

The base sequence of insert DNA of the clone described above was determined using primers (Pharmacia, M13 Universal Primer (SEQ ID NO: 7, 5`-fluorescein-**cgacgttgtaaaacgacggccagt**-3`) and M13 Reverse Primer (SEQ ID NO: 8, 5`-fluorescein-**caggaaacagctatgac**-3`)).

[0024]

From this base sequence, an open reading frame was selected by matching it to the collectin amino acid sequence, and the base sequence corresponding to the amino acid sequence which could be read from the open reading frame was extracted, then, primers for digoxigenin (DIG) labeled cDNA probes (Reverse Primer: **caatctgatgagaaggtgatg** (SEQ ID NO: 4) and Forward Primer: **acgaggggctggatgggacat** (SEQ ID NO: 5)) corresponding to the parts of the base sequences were produced using DNA/RNA Synthesizer of Applied Biosystems, 392A. DIG labeling was performed using PCR DIG Probe Synthesis Kit (Boehringer Mannheim). The reaction mixture contained: plasmid DNA (clone W72977, 50 ng/ 1), 2 1 (100 ng); 10 x Buffer 5 1; 25 mM $MgCl_2$, 5 1; dNTP (PCR Labeling Mix), 5 1; 20 M Reverse Primer, 2.5 1; 20 M Forward Primer, 5 1; $H_2O$, 28 1; Taq Polymerase, 0.5 1. PCR reaction was performed with Zymoreactor of ATTO Corp. through 35 cycles of: 1 minute at 92 C, 1 minute at 55 C, and 2 minutes at 72 C.

[0025]

Example 3: Screening of cDNA Library from Human Placenta

First, phage cDNA library from human placenta was titrated as follows. Escherichia coli Y1090r⁻, 0.2 ml, which had been cultured at 37 C for 16 hours in mLB medium (LB medium (1 g tryptone, 0.5g yeast extract and 0.5 g NaCl in total volume of 100 ml) containing 10 mM $MgSO_4$ and 0.2% maltose) and 0.1 ml of

cDNA library serially diluted with SM buffer (5.8 g NaCl, 2 g $MgSO_4 \cdot 7H_2O$, 2 M Tris-HCl (pH 7.5) 25 ml, and 2% gelatin 5 ml in total volume of 1L) were incubated at 37 C for 15 minutes, then the mixtures were added to 2.5 ml of LB-TOP agarose (0.75% agarose/LB medium) to make homogenous solutions, and plated onto 90 mm$\phi$ LB Medium Plates (Iwaki Glass), (1.5% agar/LB medium). The added solutions were hardened at room temperature for 15 minutes, then incubated for 5 hours at 42 C. The plaques on each of the plates were counted, and the titer of the phage was calculated. Consequently, the titer calculated to be $2.1 \times 10^{10}$ pfu/ml. The screening was performed as follows using the probe prepared in Example 2.

【0026】

Escherichia coli Y1090r$^-$, 0.6 ml, which had been cultured at 37 C for 16 hours in mLB medium, and cDNA library diluted with SM buffer to $1 \times 10^5$ pfu were incubated at 37 C for 15 minutes, then the mixture was added to 7.5 ml of LB-TOP agarose (0.75% agarose) to make a homogenous solution. The solution was plated onto ten LB square plates of 140 $cm^2$ (Nissui Seiyaku), hardened at room temperature for 15 minutes, then the plates were incubated for 5 hours at 42 C. After plaque formation of each of the plates was confirmed, the transfer to the nylon membranes was performed. The transfer was performed using Nytran 13N (Schleicher and Schuell Co.). The filters (12.5 cm x 9.0 cm in size) were immersed in distilled water for 10 minutes to be wet, then the excess water was removed on Whatmann 3MM Paper, and the filters were placed on the plates having the plaques formed thereon. After standing for two minutes, the filters were recovered and air-dried for 10 minutes. The phage DNA on the filters was denatured for 2 minutes with 0.2 M NaOH/1.5 M NaCl, followed by neutralization with 0.4 M Tris-HCl (pH 7.6) / 2 x SSC for 2 minutes and washing with 2 x SSC for 2 minutes. Thereafter, the phage DNA was fixed on the membrane by UV irradiation with GS GENE LINKER (BioRad). Hybridization, and detection of the signals were conducted as follows. The filters were soaked in 2 x SSC, and the excess moisture was removed using

Whatmann 3MM Paper, then the filters were placed in a hybridization bag and prehybridization at 68 C for one hour in a hybridization solution (5 x SSC, 1% blocking agent, 0.1% N-lauroyl sarcosine and 0.02% SDS) was performed. Subsequently, the hybridization solution was removed from the bag, and the hybridization solution containing DIG labeled cDNA probe at a concentration of 10 ng/ml was added thereto, and hybridization was proceeded at 55 C for 16 hours. After the hybridization was completed, the filters were washed twice in a solution of 2 x SSC/0.1% SDS for 5 minutes; and further washed twice in a solution of 0.5 x SSC/0.1% SDS for 15 minutes. Then, SDS was removed using DIG buffer I (100 mM Tris-HCl, 150 mM NaCl (pH 7.5)) for 1 minute, and the filters were blocked with DIG buffer II (1% blocking agent in DIG buffer I) for 30 minutes. After washing the filters with DIG buffer I for one minute, a solution of alkaline phosphatase labeled anti-DIG antibody (Boehringer Mannheim) which was diluted to 5000-fold in DIG buffer II was added, and the reaction between antigen and antibody were allowed for 30 minutes at room temperature, then the filters were washed twice with DIG buffer I for 15 minutes at room temperature. Through the subsequent treatment of the filters with DIG buffer III (100 mM Tris-HCl, 100 mM NaCl (pH 9.5), 50 mM $MgCl_2$) for 3 minutes, the concentration of $Mg^{2+}$ was elevated, when a solution of NBT/BCIP (WAKO Chem., Co.) in DIG buffer III was added for color development, 10 positive clones were identified. The plaques corresponding to these clones were excised from the plates and placed in the tubes containing 1 ml of SM buffer, after stirring for 10 minutes, each of the buffer solution was serially diluted with SM buffer, and 0.1 ml of the diluted solution was mixed with 0.2 ml cultures of Escherichia coli Y1090r$^-$ which had been cultured in mLB medium for 16 hours at 37 C, thereafter, the mixture was incubated for 15 minutes at 37 C. Then the mixed solution was added to 2.5 ml of LB-TOP agarose to make a homogenous solution, the solution was plated into ten 90mm$\phi$ LB plates, hardened at room temperature for 15 minutes, then the plates were incubated for 5 hours at 42 C, thereby, several

plaques were obtained, and the secondary screening was performed essentially in accordance with the procedures of the primary screening.

【0027】

Example 4: Sequencing of Novel Collectin Base Sequence

The plaque of the clone that was expected as being appropriate among the positive clones obtained in the above secondary screening was excised from the plates, then was placed into a tube containing distilled water 200 l followed by stirring for 30 minutes at room temperature, and the tube was centrifuged at 15,000 rpm for 5 minutes, and the supernatant was obtained therefrom.

【0028】

The insert DNA was amplified by PCR with TaKaRa LA PCR Kit Ver.2 (TAKARA Syuzo, Co.) using the resulting supernatant as a template. PCR reactions contained: the supernatant, 27 l; 10 x LA PCR Buffer II ($Mg^{2+}$ free), 5 l; 25 mM $MgCl_2$, 5 l; dNTP Mix, 8 l; 20 M gt11 Reverse Primer (SEQ ID NO: 9: 5'-**ttgacaccagaccaactggtaatg**-3'), 2.5 l; 20 M gt11 Forward Primer (SEQ ID NO: 10: 5'-**ggtggcgacgactcctggagcccg**-3'), 2.5 l; LA Taq polymerase, 0.5 l; and $H_2O$, to make final volume of 50 l. PCR reaction was performed using Applied Biosystems Gene Amp PCR System 9600, with 30 cycles of: 20 seconds at 98 C, and 5 minutes at 68 C. The PCR product was verified by the electrophoresis on 1% agarose gel, and purified through excising from the gel. For this purification step, Sephaglas BandPrep Kit (Pharmacia) was used.

【0029】

The excised DNA fragment was incorporated into pCR2.1 vector (Invitrogen, TA Cloning Kit). The recombinant vector was transformed into TOP10F' cell included in the Invitrogen TA Cloning Kit. The transformants were cultured in LB medium (containing 100 g/ml ampicillin), and three kinds of plasmids were extracted by alkaline SDS method.

[0030]

Thus obtained DNA was cleaved with restriction enzymes that were expected to be adequate, and each DNA fragment was incorporated into pUC18 vector followed by transformation into XL1-Blue cell. The transformants were cultured on LB medium (containing 100 g/ml ampicillin), and the plasmids were extracted by alkaline SDS method. CL-P1-2-1 resulted in a plasmid containing EcoRI-Hind III fragment and Hind III-EcoRI fragment; CL-P1-3-4 resulted in a plasmid containing EcoRI-BamHI fragment, BamHI-SmaI fragment, SmaI-HindIII fragment, KpnI-Sau3AI fragment, Sau3AI-EcoRI fragment, EcoRI-KpnI fragment and EcoRI-SmaI fragment; CL-P1-3-7 resulted in a plasmid containing EcoRI-BamHI fragment, BamHI-SmaI fragment, SmaI-HindIII fragment, KpnI-Sau3AI fragment, Sau3AI-EcoRI fragment, EcoRI-KpnI fragment and KpnI-EcoRI fragment. The primers were prepared by synthesizing with DNA/RNA synthesizer the following primers labeled with M13 Universal Primer (SEQ ID NO: 5) and M13 Reverse Primer (SEQ ID NO: 6) respectively attached to the Autoread Sequencing Kit as well as FITC (Pharmacia, Fluore Prime), and were their entire regions were base sequenced with Autoread Sequencing Kit (Pharmacia) and A.L.F. Autosequencer.

[0031]

HPP 1: 5'-fluorescein-**cgtgaaaatgaatggaagtgg**-3' (SEQ ID NO: 11),
HPP 2: 5'-fluorescein-**ttttatccattgctgttcctc**-3' (SEQ ID NO: 12),
HPP 3: 5'-fluorescein-**ctggcagtccccgaggtccag**-3' (SEQ ID NO: 13),
HPP 5: 5'-fluorescein-**gctggtccccccggagagcgt**-3' (SEQ ID NO: 14)

The outline of the sequencing strategy performed is shown in Figure 4. An ORF of the obtained collectin is illustrated in Figure 4 (a), wherein a collagen-like region is denoted as G-X-Y. Further, in Figure 4 (b), name of each primer and positions of the base sequences determined by the sequencer (shown as allows), and M13 Universal Primer (shown as U) as well as M13 Reverse Primer (shown as R) are illustrated.

[0032]

Further, a base sequence around the 5'-end region comprising a transcription initiation site was determined using Cap site cDNA.

[0033]

First PCR was performed with Cap Site cDNA, on Human Liver (NIPPON GENE) using TGP1 Primer (5'-**tcttcagtttccctaatccc**-3' (SEQ ID NO: 16)) that was synthesized with the attached 1RC2 Primer (5'-**caaggtacgccacagcgtatg**-3' (SEQ ID NO: 15)) and 392A DNA/RNA Synthesizer (Applied Biosystems). The employed reaction mix solution contained LA PCR Buffer II ($Mg^{2+}$ free), 2.5 mM $MgCl_2$, each 200 M of dATP, dCTP, dGTP and dTTP (all of which are manufactured by TAKARA Syuzo, Co.), 1 l; Cap Site cDNA Human Liver; 0.5 $\mu$ 1RC2 Primer (both of which are manufactured by NIPPON GENE), and 0.5 M TGP1 Primer, in total volume of 50 l. PCR was performed using a program comprising 35 cycles of: heat denaturation for 20 seconds at 95 C, anealing for 20 seconds at 60 C, extension for 20 seconds at 72 C, with heat denaturation for 5 minutes at 95 C prior to the repeated reaction and final extension for 10 minutes at 72 C. After completing the first PCR, nested PCR was conducted. The reaction was performed using 1 l of the first PCR product as a template, together with primers of 2RC2 Primer (5'-**gtacgccaca gcgtatgatgc**-3' (SEQ ID NO: 17)) attached and of synthetic TGP2 Primer (5'-**cattcttgacaaacttcatag**-3' (SEQ ID NO: 18)) that was synthesized similarly to TGP1 Primer, and with the same reaction components and program (except that the cycle number was 25) as in the first PCR. The PCR reaction was performed with TaKaRa PCR Thermal Cycler 480. After thus obtained PCR product was confirmed on agarose gel electrophoresis, the band was excised from the gel, followed by freezing at -80 C for 10 minutes, centrifuge at 15000 rpm, for 10 minutes, and then the supernatant was purified by ethanol precipitation.

[0034]

The purified DNA fragment was incorporated into pT7Blue Vector (Novagen), and the vector was transformed into competent

XL1-Blue cell. The transformants were cultured on LB medium (containing 100 g/ml ampicillin), and the plasmids were extracted by alkaline SDS method, followed by sequencing of the base sequence with Autoread Sequencing Kit (Pharmacia) and A.L.F. DNA Sequencer. The employed primers were M13 Universal Primer (SEQ ID NO: 7) and M13 Reverse Primer (SEQ ID NO: 8) attached to AutoRead Sequencing Kit.

[0035]

As a result thereof, it was confirmed that the cDNA clone of the novel collectin that was obtained in Example 3 contained 2024 bases comprising ORF (open reading frame) of 1026 bases encoding 342 amino acids set out in SEQ ID NO: 2.

[0036]

Next, when the homology of the DNA and amino acid sequence was searched on GenBank database, the results revealed that the amino acid sequence of the obtained protein is distinct from those of the collectins identified previously and is therefore derived from a novel protein.

[0037]

In addition, the amino acid sequence of the novel collectin of the present invention was compared to those of three collectin proteins reported in the art. The alignment is shown in Figures 5 and 6. Similarly to Figures 2 and 3, homologous amino acid residues were boxed. This alignment suggests that the obtained novel protein shares homology with known collectins and it belongs to the collectin family.

[0038]

Example 5: Genomic Southern Analysis of Novel Collectin

Genomic Southern analysis was performed in order to clarify as to whether the novel collectin gene comprising the cDNA sequence verified in Example 4 was a single copy gene or a multi copy gene.

[0039]

Four g of human genomic DNA (Promega) from human blood was digested with any of the restriction enzymes of (1) EcoRI, (2) XbaI, (3) HindIII, (4) PstI, (5) BglII or (6) BamHI, followed

by electrophoresis on 0.8% agarose gel at 100 mA, for 3 hours. After the electrophoresis was completed, they were transferred to a nylon membrane (Nytran 13N) to prepare a membrane for the analysis. For the transfer step, the electrophoresed gel was first immersed in 100 ml of 0.25 N HCl for 10 minutes, washed three times with distilled water, then immersed twice in 100 ml of a denaturalizing solution (1.5 M NaCl, 0.5 M NaOH) for 15 minutes, and immersed in 100 ml of a neutralizing solution (0.5 M Tris-HCl, 3 M NaCl (pH 6.8)) for 30 minutes so that depurination, denaturation and neutralization were accomplished, the DNA was then transferred using Vacuum Blotting System (Toyobo Engineering, VB-30). In this step, the membrane is used which had been pretreated by immersing it in 2 x SSC for 5 minutes and in 20 x SSC for 5 minutes, while a pad is used which had been soaked with 20 x SSC. After the transfer was terminated, fixation of the DNA was performed by UV irradiation.

[0040]

As a hybridization probe employed for the Southern analysis, DNA probe consisting of the base sequence of: **gaagacaagt cttcacatct tgttttcata aacactagag aggaacagca atggataaaa aaacagatgg tagggagaga gagccactgg atcggcctca cagactcaga g** (SEQ ID NO: 21) was used, wherein it was prepared by labeling a portion of ORF in the cDNA sequence of the novel collectin according to Example 4 with primers:

5'-gaagacaagtcttcaactcttg-3' (SEQ ID NO: 19),

5'-ctctgagtctgtgaggccgatc-3' (SEQ ID NO: 20), and

the above-mentioned PCR DIG Probe Synthesis Kit.

Prior to the hybridization, the probe was boiled for 10 minutes, and was rapidly frozen with dry ice/ethanol for 5 minutes.

[0041]

First of all, membrane treated for a transcription was immersed in 2 x SSC for 5 minutes, then prehybridization was performed in 10ml of ExpressHyb Hybridization Solution (Clonetech) at 65 C for 30 minutes. Subsequently, the above frozen probe was diluted to 10 ng/ml in ExpressHyb Hybridization Solution, and 2 ml of this solution was used for hybridization

at 65 C for one hour.

[0042]

Following the hybridization, the membrane was washed by: shaking twice in 20 ml of 2 x SSC, 0.1% SDS solution at room temperature for 5 minutes, then twice in 20 ml of 0.2 x SSC, 0.1% SDS solution at 65 C for 15 minutes. Next, the membrane was washed twice with 50 ml of DIG buffer I (100 mM Tris-HCl, 150 mM NaCl (pH 7.5)) at room temperature for one minute in order to remove SDS, and was blocked in 50 ml of DIG buffer II' (1.5% blocking agent, DIG buffer I) at room temperature for one hour. Thereafter, the membrane was treated for 30 minutes with 10 ml of alkaline phosphatase labeled anti-DIG antibody which had been diluted to 5000-fold in DIG buffer I containing 0.2% Tween20 followed by washing twice by shaking in 50 ml of DIG buffer I which contains 0.2% Tween20 at room temperature for 20 minutes. After soaking the membrane twice in 10 ml of DIG buffer III at room temperature for 3 minutes, it was placed in a hybridization bag, and CSPD (registered trade name, Boehringer Mannheim: chemiluminescence substrate) that had been diluted to 100-fold in DIG buffer III was added thereto so that the solution can spread over the membrane. Subsequently, the membrane was exposed onto Instant Film T612 (Polaroid).

[0043]

Consequently, it was speculated that the gene of the obtained novel collectin has been a single copy gene, because only one or two signal/s could be detected from the respective genomic DNA which was digested with each of the restriction enzymes, as shown in the lanes of Figure 7.

[0044]

Example 6: Analysis of Expression Distribution in Human Tissue by Novel Collectin

In order to examine the expression of the mRNA of the novel collectin of the present invention in various human tissues, analysis was performed by RT-PCR.

[0045]

RT-PCR was performed using RNA LA PCR Kit (AMV) Ver.1.1 (TAKARA Syuzo, Co.) with each RNA taken from several human tissues ((1) brain, (2) heart, (3) kidney, (4) spleen, (5) liver, (6) small intestine, (7) muscle, (8) testis, (9) placenta, or (10) colon (OriGene Technologies, Inc.)) as a template. First, reverse transcription reaction was conducted in the following reaction mixture. The reaction mixture contained 5 mM $MgCl_2$; 1 x RNA PCR Buffer, 1 mM dNTP Mixture, 1 U/ 1 Rnase inhibitor, 0.25 U/ 1 reverse transcriptase, 0.125 M Oligo dT-Adaptor Primer, RNA 1 g, and was adjusted to total volume of 20 1 with RNase free distilled water. At the same time, a reaction mixture without reverse transcriptase was also prepared for the negative control. The reaction solution was placed in 0.2 ml tube, and subjected to PCR with TaKaRa PCR Thermal Cycler PERSONAL (TAKARA Syuzo, Co.) through 1 cycle of: 30 minutes at 42 C, 5 minutes at 99 C, and 5 minutes at 5 C. Thus resulted PCR product was subsequently used for LA PCR with the following reaction mixture. 2.5 mM $MgCl_2$, 1 x LA PCR Buffer II ($Mg^{2+}$ free), 2U TaKaRa LA Taq and two kinds of 0.2 M primers (RT-PCR Primer U: 5'-gtgcccctggccctgcagaatg-3' (SEQ ID NO: 22) and RT-PCR Primer R: 5'-gcatatcaccctggggaacattttag-3' (SEQ ID NO: 23) that could amplify a cDNA sequence spanning from neck region to carbohydrate recognition domain of the novel collectin are mixed and the mixture was adjusted to total volume of 80 1 with sterile distilled water. PCR was performed through 1 cycle of 2 minutes at 94 C and 50 cycles of: 30 seconds at 94 C, 30 seconds at 60 C and 90 seconds at 72 C. The reaction product was separated on 1% agarose gel electrophoresis, followed by staining with ethidium bromide solution (0.1 g/ml), verification of the electrophoretic pattern with transilluminator, and the expressed tissues were identified.

[0046]

Further, in order to compare the expressed amount in each of the tissues, RT-PCR was performed to amplify a part of -actin in each of the tissues, and the amount of RNA was corrected.

The RT-PCR was performed similarly to the above procedure with reverse transcriptase reaction, PCR reaction, and identified using 1% agarose gel electrophoresis as described above. The reaction mixture of the reverse transcription contained 5 mM $MgCl_2$, 1 x RNA PCR Buffer, 1 mM dNTP Mixture, 1U/ 1 RNase inhibitor, 0.25 U/ 1 reverse transcriptase, 2.5 M random 9 mer, RNA 10 ng, which was then adjusted to total volume of 60 1 with RNase free distilled water. PCR was performed through 1 cycle of: 10 minutes at 30 C, 15 minutes at 42 C, 5 minutes at 99 C and 5 minutes at 5 C. Thus resulted PCR product was subsequently used for PCR with the following reaction mixture. 2.5 mM $MgCl_2$, 1 x LA PCR Buffer II ($Mg^{2+}$ free), 2U TaKaRa LA Taq and 0.25 M human -actin sense primer 5'-**caagagatggccacggctgct**-3'(SEQ ID NO: 24), 0.25 M human -actin antisense primer 5'-**tccttctgcatcctgt cggca**-3'(SEQ ID NO: 25) are mixed and the mixture was adjusted to total volume of 40 1 with sterile distilled water. PCR was performed through 30 cycles of: 15 seconds at 94 C, and 30 seconds at 68 C.

[0047]

The results are shown in Figure 8, suggesting that mRNA of the novel collectin of the present invention has been expressed in placenta (lane 9), spleen (lane 4), and kidney (lane 3), but extremely high expression in placenta is clearly suggested.

[0048]

Example 7: Genomic Southern Analysis of Novel Collectin from Various Animals

In order to elucidate conservation of the collectin gene of the present invention in the other species of animals, analysis by genomic Southern hybridization was performed.

[0049]

As a hybridization probe, DNA probe labeled with DIG prepared by labeling, with the above-described PCR DIG Probe Synthesis Kit (Boehringer Mannheim), a portion corresponding to ORF in the cDNA sequence of the novel collectin as described above was used, while the employed membranes were prepared by

treating, with restriction enzyme EcoRI, 5 g of each genomic DNAs of (1) human (Promega), (2) monkey (Clonetech), (3) rat (Promega), (4) mouse (Promega), (5) dog (Clonetech), (6) cow (Promega), (7) rabbit (Clonetech), and (8) chicken (Promega), electrophoresisingh the DNAs on agarose gel, transferring them to Nytran 13N membrane and fixing the same by UV irradiation.

[0050]

Using such probe and membrane, hybridization was performed according to the following procedures. First, the membrane was immersed in 2 x SSC for 5 minutes, then prehybridization was performed in 10 ml of ExpressHyb Hybridization Solution at 65 C for 30 minutes. Subsequently, the probe that had been frozen as described above was diluted in the ExpressHyb Hybridization Solution to be 10 ng/ml, and 2 ml of thus diluted probe solution was used for hybridization at 65 C for one hour.

[0051]

Following the hybridization, the membrane was washed by: shaking twice in 20 ml of 2 x SSC, 0.1% SDS at room temperature for 5 minutes, and then shaking twice in 20 ml of 0.2 x SSC, 0.1% SDS at 68 C for 15 minutes. Next, the membrane was washed twice with DIG buffer I at room temperature for one minute in order to remove SDS, and was blocked in 50 ml of DIG buffer II' at room temperature for one hour. Thereafter, the membrane was treated for 30 minutes with 10 ml of alkaline phosphatase labeled anti-DIG antibody which had been diluted to 5000-fold in DIG buffer I which contains 0.2% Tween20 followed by washing twice with shaking in 50 ml of DIG buffer I containing 0.2% Tween20 at room temperature for 20 minutes. After soaking the membrane twice in 10 ml of DIG buffer III at room temperature for 3 minutes, it was placed in a hybridization bag, and CSPD diluted to 100-fold in DIG buffer III was added thereto so that the solution can spread over the membrane, the membrane was subsequently exposed to Instant Film T612.

[0052]

The result of this analysis is shown in Figure 9, wherein clear signals can been found in all lanes except for the lane on chicken (Lane 8), it was therefore demonstrated that the novel collectin gene of the present invention has been conserved between the mammalian species.

[0053]

Example 8: Genetic Analysis of Novel Collectin

To elucidate the genetic positional relevance of the present collectin against the known collectins, analysis was performed based on the DNA sequence of the novel collectin as obtained, and a phylogenetic tree was created.

[0054]

The collectins selected as subjects for analysis were several kinds of proteins belonging to the collectin family shown in Figure 10 (in Figure, the novel collectin of the present invention is denoted as CL-P1, while a collectin from human liver which was recently isolated by the present inventor is denoted as CL-L1 (See, the specification of the Japanese Patent Application No. Hei 10-11281)), then multiple alignments were produced by clustalw method using the regions containing lectin domains based on the data obtained by searching each amino acid sequence on GenBank database, and the phylogenetic tree was created based on such alignments by N-J (neighbor-joining) method using Phylip Version 3.57c package program.

[0055]

Consequently, as shown in Figure 10, although SP-D, bovine CL-43 and bovine conglutinin have constituted single cluster, additionally, MBP and SP-A have respectively constituted separate clusters, while the collectin gene of the present invention has not belonged to any of these clusters similarly to CL-L1. Furthermore, it was speculated that the collectin of the present invention may constitute a distinct cluster which is genetically distinguishable from those of the conventional collectins including CL-L1.

【0056】

【EFFECTS OF THE INVENTION】

As stated above, the present invention provides the novel collectin gene as well as the novel collectin protein that have the characteristic structures to be seen in the known collectin and are different from the previously reported collectin.

【0057】

【SEQUENCE LISTING】

SEQUENCE LISTING

<110> FUSO PHARMACEUTICAL INDUSTRIES, LTD.

<120> Novel Collectin

<130> 98PA0270

<160> 23

<210> 1

<211> 2024

<212> DNA

<213> Homo Sapiens

<220>

<221> CDS

<222> (670)..(1695)

<400> 1

gtcacgaatc tgcagcaaga taccagcgtg ctccagggca atctgcagaa ccaaatgtat 60

tctcataatg tggtcatcat gaacctcaac aacctgaacc tgacccaggt gcagcagagg 120

aacctcatca cgaatctgca gcggtctgtg gatgacacaa gccaggctat ccagcgaatc 180

aagaacgact ttcaaaatct gcagcaggtt tttcttcaag ccaagaagga cacggattgg 240

ctgaaggaga aagtgcagag cttgcagacg ctggctgcca acaactctgc gttggccaaa 300

gccaacaacg acaccctgga ggatatgaac agccagctca actcattcac aggtcagatg 360

gagaacatca ccactatctc tcaagccaac gagcagaacc tgaaagacct gcaggactta 420

cacaaagatg cagagaatag aacagccatc aagttcaacc aactggagga acgcttccag 480

ctctttgaga cggatattgt gaacatcatt agcaatatca gttacacagc ccaccacctg 540

cggacgctga ccagcaatct aaatgaagtc aggaccactt gcacagatac ccttaccaaa 600

cacacagatg atctgacctc cttgaataat accctggcca acatccgttt ggattctgtt 660

tctctcagg atg caa caa gat ttg atg agg tcg agg tta gac act gaa gta 711

```
                Met Gln Gln Asp Leu Met Arg Ser Arg Leu Asp Thr Glu Val
                  1               5                  10
gcc aac tta tca gtg att atg gaa gaa atg aag cta gta gac tcc aag     759
Ala Asn Leu Ser Val Ile Met Glu Glu Met Lys Leu Val Asp Ser Lys
 15                  20                  25                  30
cat ggt cag ctc atc aag aat ttt aca ata cta caa ggt cca ccg ggc     807
His Gly Gln Leu Ile Lys Asn Phe Thr Ile Leu Gln Gly Pro Pro Gly
                35                  40                  45
ccc agg ggt cca aga ggt gac aga gga tcc cag gga ccc cct ggc cca     855
Pro Arg Gly Pro Arg Gly Asp Arg Gly Ser Gln Gly Pro Pro Gly Pro
            50                  55                  60
act ggc aac aag gga cag aaa gga gag aag ggg gag cct gga cca cct     903
Thr Gly Asn Lys Gly Gln Lys Gly Glu Lys Gly Glu Pro Gly Pro Pro
        65                  70                  75
ggc cct gcg ggt gag aga ggc cca att gga cca gct ggt ccc ccc gga     951
Gly Pro Ala Gly Glu Arg Gly Pro Ile Gly Pro Ala Gly Pro Pro Gly
    80                  85                  90
gag cgt ggc ggc aaa gga tct aaa ggc tcc cag ggc ccc aaa ggc tcc     999
Glu Arg Gly Gly Lys Gly Ser Lys Gly Ser Gln Gly Pro Lys Gly Ser
 95                  100                 105                 110
cgt ggt tcc cct ggg aag ccc ggc cct cag ggc ccc agt ggg gac cca    1047
Arg Gly Ser Pro Gly Lys Pro Gly Pro Gln Gly Pro Ser Gly Asp Pro
                115                 120                 125
ggc ccc ccg ggc cca cca ggc aaa gag gga ctc ccc ggc cct cag ggc    1095
Gly Pro Pro Gly Pro Pro Gly Lys Glu Gly Leu Pro Gly Pro Gln Gly
            130                 135                 140
cct cct ggc ttc cag gga ctt cag ggc acc gtt ggg gag cct ggg gtg    1143
Pro Pro Gly Phe Gln Gly Leu Gln Gly Thr Val Gly Glu Pro Gly Val
        145                 150                 155
cct gga cct cgg gga ctg cca ggc ttg cct ggg gta cca ggc atg cca    1191
Pro Gly Pro Arg Gly Leu Pro Gly Leu Pro Gly Val Pro Gly Met Pro
    160                 165                 170
ggc ccc aag ggc ccc ccc ggc cct cct ggc cca tca gga gcg gtg gtg    1239
Gly Pro Lys Gly Pro Pro Gly Pro Pro Gly Pro Ser Gly Ala Val Val
175                 180                 185                 190
ccc ctg gcc ctg cag aat gag cca acc ccg gca ccg gag gac aat ggc    1287
```

```
Pro Leu Ala Leu Gln Asn Glu Pro Thr Pro Ala Pro Glu Asp Asn Gly
                195                 200                 205
tgc ccg cct cac tgg aag aac ttc aca gac aaa tgc tac tat ttt tca   1335
Cys Pro Pro His Trp Lys Asn Phe Thr Asp Lys Cys Tyr Tyr Phe Ser
            210                 215                 220
gtt gag aaa gaa att ttt gag gat gca aag ctt ttc tgt gaa gac aag   1383
Val Glu Lys Glu Ile Phe Glu Asp Ala Lys Leu Phe Cys Glu Asp Lys
        225                 230                 235
tct tca cat ctt gtt ttc ata aac act aga gag gaa cag caa tgg ata   1431
Ser Ser His Leu Val Phe Ile Asn Thr Arg Glu Glu Gln Gln Trp Ile
    240                 245                 250
aaa aaa cag atg gta ggg aga gag agc cac tgg atc ggc ctc aca gac   1479
Lys Lys Gln Met Val Gly Arg Glu Ser His Trp Ile Gly Leu Thr Asp
255                 260                 265                 270
tca gag cgt gaa aat gaa tgg aag tgg ctg gat ggg aca tct cca gac   1527
Ser Glu Arg Glu Asn Glu Trp Lys Trp Leu Asp Gly Thr Ser Pro Asp
                275                 280                 285
tac aaa aat tgg aaa gct gga cag ccg gat aac tgg ggt cat ggc cat   1575
Tyr Lys Asn Trp Lys Ala Gly Gln Pro Asp Asn Trp Gly His Gly His
            290                 295                 300
ggg cca gga gaa gac tgt gct ggg ttg att tat gct ggg cag tgg aac   1623
Gly Pro Gly Glu Asp Cys Ala Gly Leu Ile Tyr Ala Gly Gln Trp Asn
        305                 310                 315
gat ttc caa tgt gaa gac gtc aat aac ttc att tgc gaa aaa gac agg   1671
Asp Phe Gln Cys Glu Asp Val Asn Asn Phe Ile Cys Glu Lys Asp Arg
    320                 325                 330
gag aca gta ctg tca tct gca tta taacggactg tgatgggatc acatgagcaa   1725
Glu Thr Val Leu Ser Ser Ala Leu
335                 340
attttcagct ctcaaaggca aaggacactc ctttctaatt gcatcacctt ctcatcagat  1785
tgaaaaaaaa aaaagcactg aaaaccaatt actgaaaaaa aattgacagc tagtgttttt  1845
taccatccgt cattacccaa agacttggga actaaaatgt tccccagggt gatatgctga  1905
ttttcattgt gcacatggac tgaatcacat agattctcct ccgtcagtaa ccgtgcgatt  1965
atacaaatta tgtcttccaa agtatggaac actccaatca gaaaaaggtt atcatcccg   2024
```

<210> 2

<211> 547

<212> PRT

<213> Homo Sapiens

<220>

<223> Deduced Amino Acid Sequence of Novel Collectin from Nucleotide Sequence.

<400> 2

```
Met Tyr Ser His Asn Val Val Ile Met Asn Leu Asn Asn Leu Asn Leu
 1               5                  10                  15
Thr Gln Val Gln Gln Arg Asn Leu Ile Thr Asn Leu Gln Arg Ser Val
            20                  25                  30
Asp Asp Thr Ser Gln Ala Ile Gln Arg Ile Lys Asn Asp Phe Gln Asn
        35                  40                  45
Leu Gln Gln Val Phe Leu Gln Ala Lys Lys Asp Thr Asp Trp Leu Lys
    50                  55                  60
Glu Lys Val Gln Ser Leu Gln Thr Leu Ala Ala Asn Asn Ser Ala Leu
65                  70                  75                  80
Ala Lys Ala Asn Asn Asp Thr Leu Glu Asp Met Asn Ser Gln Leu Asn
                85                  90                  95
Ser Phe Thr Gly Gln Met Glu Asn Ile Thr Thr Ile Ser Gln Ala Asn
            100                 105                 110
Glu Gln Asn Leu Lys Asp Leu Gln Asp Leu His Lys Asp Ala Glu Asn
        115                 120                 125
Arg Thr Ala Ile Lys Phe Asn Gln Leu Glu Glu Arg Phe Gln Leu Phe
    130                 135                 140
Glu Thr Asp Ile Val Asn Ile Ile Ser Asn Ile Ser Tyr Thr Ala His
145                 150                 155                 160
His Leu Arg Thr Leu Thr Ser Asn Leu Asn Glu Val Arg Thr Thr Cys
                165                 170                 175
Thr Asp Thr Leu Thr Lys His Thr Asp Asp Leu Thr Ser Leu Asn Asn
            180                 185                 190
Thr Leu Ala Asn Ile Arg Leu Asp Ser Val Ser Leu Arg Met Gln Gln
        195                 200                 205
Asp Leu Met Arg Ser Arg Leu Asp Thr Glu Val Ala Asn Leu Ser Val
    210                 215                 220
Ile Met Glu Glu Met Lys Leu Val Asp Ser Lys His Gly Gln Leu Ile
```

225 230 235 240

Lys Asn Phe Thr Ile Leu Gln Gly Pro Pro Gly Pro Arg Gly Pro Arg

245 250 255

Gly Asp Arg Gly Ser Gln Gly Pro Pro Gly Pro Thr Gly Asn Lys Gly

260 265 270

Gln Lys Gly Glu Lys Gly Glu Pro Gly Pro Pro Gly Pro Ala Gly Glu

275 280 285

Arg Gly Pro Ile Gly Pro Ala Gly Pro Pro Gly Glu Arg Gly Gly Lys

290 295 300

Gly Ser Lys Gly Ser Gln Gly Pro Lys Gly Ser Arg Gly Ser Pro Gly

305 310 315 320

Lys Pro Gly Pro Gln Gly Pro Ser Gly Asp Pro Gly Pro Pro Gly Pro

325 330 335

Pro Gly Lys Glu Gly Leu Pro Gly Pro Gln Gly Pro Pro Gly Phe Gln

340 345 350

Gly Leu Gln Gly Thr Val Gly Glu Pro Gly Val Pro Gly Pro Arg Gly

355 360 365

Leu Pro Gly Leu Pro Gly Val Pro Gly Met Pro Gly Pro Lys Gly Pro

370 375 380

Pro Gly Pro Pro Gly Pro Ser Gly Ala Val Val Pro Leu Ala Leu Gln

385 390 395 400

Asn Glu Pro Thr Pro Ala Pro Glu Asp Asn Gly Cys Pro Pro His Trp

405 410 415

Lys Asn Phe Thr Asp Lys Cys Tyr Tyr Phe Ser Val Glu Lys Glu Ile

420 425 430

Phe Glu Asp Ala Lys Leu Phe Cys Glu Asp Lys Ser Ser His Leu Val

435 440 445

Phe Ile Asn Thr Arg Glu Glu Gln Gln Trp Ile Lys Lys Gln Met Val

450 455 460

Gly Arg Glu Ser His Trp Ile Gly Leu Thr Asp Ser Glu Arg Glu Asn

465 470 475 480

Glu Trp Lys Trp Leu Asp Gly Thr Ser Pro Asp Tyr Lys Asn Trp Lys

485 490 495

Ala Gly Gln Pro Asp Asn Trp Gly His Gly His Gly Pro Gly Glu Asp

500 505 510

Cys Ala Gly Leu Ile Tyr Ala Gly Gln Trp Asn Asp Phe Gln Cys Glu

515 520 525

Asp Val Asn Asn Phe Ile Cys Glu Lys Asp Arg Glu Thr Val Leu Ser

530 535 540

Ser Ala Leu

545

<210> 3

<211> 27

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified Consensus Sequence of collectins Hybridizable with Novel Collectin.

<400> 3

Glu Lys Cys Val Glu Met Tyr Thr Asp Gly Lys Trp Asn Asp Arg Asn

1 5 10 15

Cys Leu Gln Ser Arg Leu Ala Ile Cys Glu Phe

20 25

<210> 4

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Reverse Primer for Screening a Novel Collectin.

<400> 4

caatctgatg agaaggtgat g 21

<210> 5

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Forward Primer for Screening a Novel Collectin.

<400> 5

acgaggggct ggatgggaca t 21

<210> 6

<211> 27

<212> PRT

<213> Artificial Sequence

<220>

<223> Consensus sequence of three collectins which were reported heretofore.

<400> 6

Glu Asp Cys Val Leu Leu Leu Lys Asn Gly Gln Trp Asn Asp Val Pro
1 5 10 15

Cys Ser Thr Ser His Leu Ala Val Cys Glu Phe
20 25

<210> 7

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> M13 Universal Primer Sequence for Sequencing.

<400> 7

cgacgttgta aaacgacggc cagt 24

<210> 8

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> M13 Reverse Primer Sequence for Sequencing.

<400> 8

caggaaaca gctatgac 17

<210> 9

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a λgt11 Reverse Primer for Sequencing.

<400> 9

ttgacaccag accaactggt aatg 24

<210> 10

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a λgt11 Forward Primer for Sequencing.

<400> 10

ggtggcgacg actcctggag cccg 24

<210> 11

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Primer for Screening a Novel Collectin.

<400> 11

cgtgaaaatg aatggaagtg g 21

<210> 12

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Primer for Screening a Novel Collectin.

<400> 12

ttttatccat tgctgttcct c 21

<210> 13

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Primer for Sequencing a Novel Collectin.

<400> 13

ctggcagtcc ccgaggtcca g 21

<210> 14

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Primer for Sequencing a Novel Collectin.

<400> 14

gctggtcccc ccggagagcg t 21

<210> 15

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a 1RC2 Primer for Cap Site Sequencing.

<400> 15

caaggtacgc cacagcgtat g 21

<210> 16

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Synthetic TGP1 Primer for Cap Site Sequencing.

<400> 16

tcttcagttt ccctaatccc 20

<210> 17

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a 2RC2 Primer for Cap Site Sequencing.

<400> 17

gtacgccaca gcgtatgatg c 21

<210> 18

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Synthetic TGP2 Primer for Cap Site Sequencing.

<400> 18

cattcttgac aaacttcata g 21

<210> 19

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Primer for Screening a Novel Collectin.

<400> 19

gaagacaagt cttcaactct tg 22

<210> 20

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Primer for Screening a Novel Collectin.

<400> 20

ctctgagtct gtgaggccga tc 22

<210> 21

<211> 111

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Probe for Screening a Novel Collectin.

<400> 21

gaagacaagt cttcacatct tgttttcata aacactagag aggaacagca atggataaaa 60

aaacagatgg tagggagaga gagccactgg atcggcctca cagactcaga g 111

<210> 22

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Forward Primer for Screening a Novel Collectin.

<400> 22

gtgcccctgg ccctgcagaa tg 22

<210> 23

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Reverse Primer for Screening a Novel Collectin.

<400> 23

gcatatcacc ctggggaaca ttttag 26

<210> 24

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Sense Primer for Screening β-Actin.

<400> 24

caagagatgg ccacggctgc t 21

<210> 25

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of an Antisense Primer for Screening β-Actin.

<400> 25

tccttctgca tcctgtcggc a 21

[BRIEF DESCRIPTION OF DRAWINGS]

[FIGURE 1] A schematic view showing basic structures and overviews of the principal collectins reported in the prior art.

[FIGURE 2] An alignment of the preceding half portions of amino acid sequences of three collectins reported in the prior art.

[FIGURE 3] An alignment of the latter half portions of the amino acid sequences as shown in Figure 2;

[FIGURE 4] Each of the primers used for sequencing the novel collectin of the present invention including (b) the nucleotide sequences which were read out from the sequencer and (a) an ORF of the obtained novel collectin;

[FIGURE 5] An alignment of the preceding half portions of amino acid sequences of the three collectins reported in the

prior art and the novel collectin of the present invention;

[FIGURE 6] An alignment of the latter half portions of the amino acid sequences in Figure 5.

[FIGURE 7] Results of genomic Southern analysis with the novel collectin of the present invention, and the restriction enzymes employed in each of the lanes are (1) EcoRI, (2) XbaI, (3) HindIII, (4) PstI, (5) BglII and (6) BamHI.

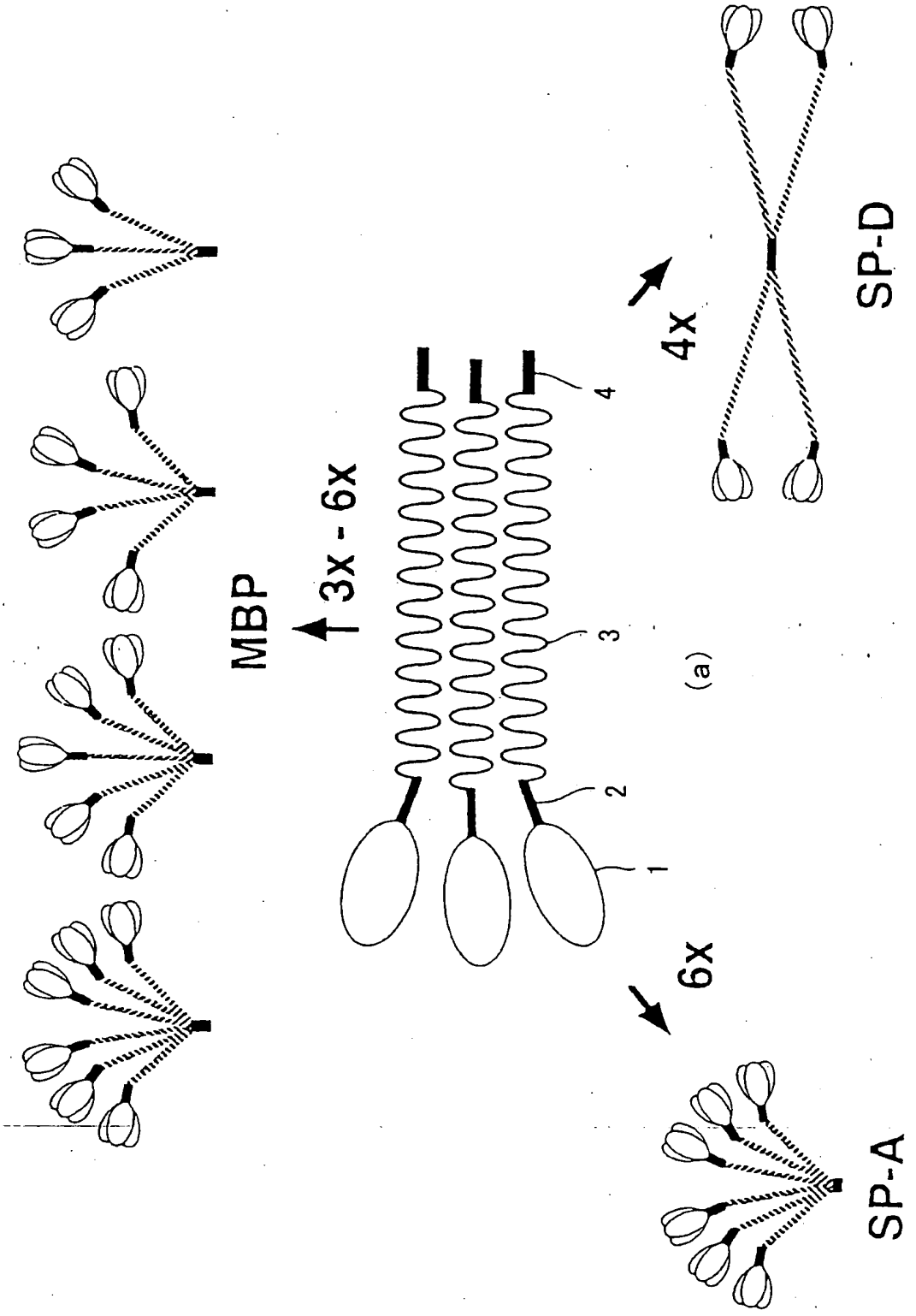
[FIGURE 8] Results of analysis of distribution of expression of mRNA in: (1) brain, (2) heart, (3) kidney, (4) spleen, (5) liver, (6) small intestine, (7) muscular tissue, (8) testis, (9) placenta, or (10) large intestine which clarify the tissue distribution of the novel collectin of the present invention.

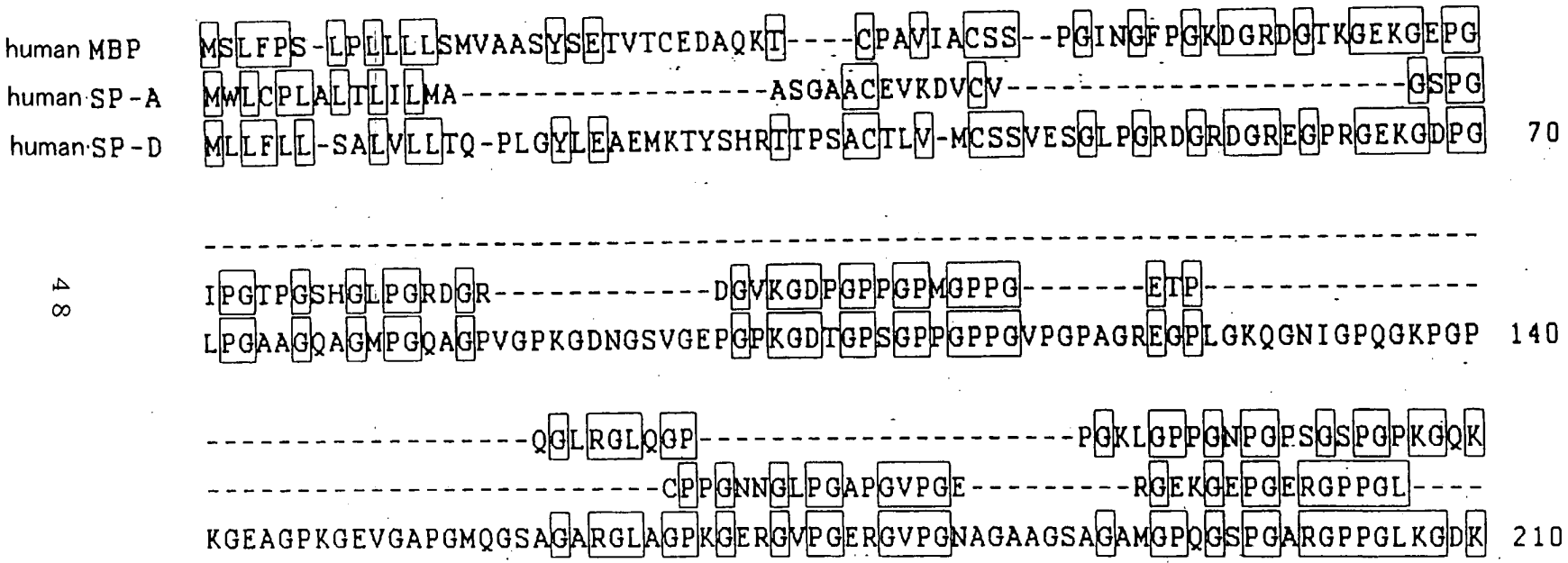
[FIGURE 9] Results of genomic Southern analysis of various vertebrates, i.e., (1) human, (2) monkey, (3) rat, (4) mouse, (5) dog, (6) cow, (7) rabbit and (8) chicken which elucidate the interspecies conservation of the novel collectin of the present invention.

[FIGURE 10] A phylogenetic tree of various collectins.

[DOCUMENT TITLE] DRAWINGS

[FIGURE 1]

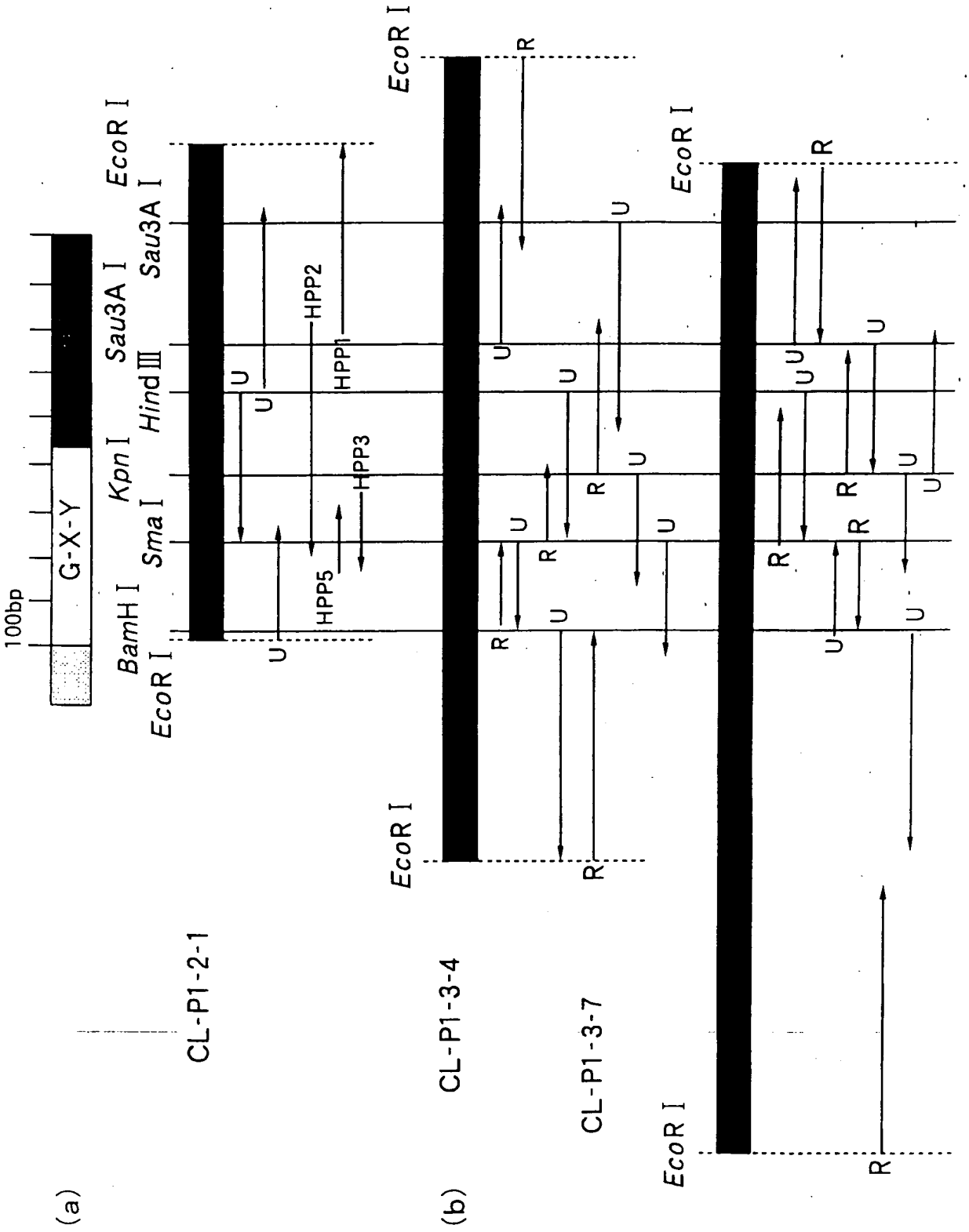


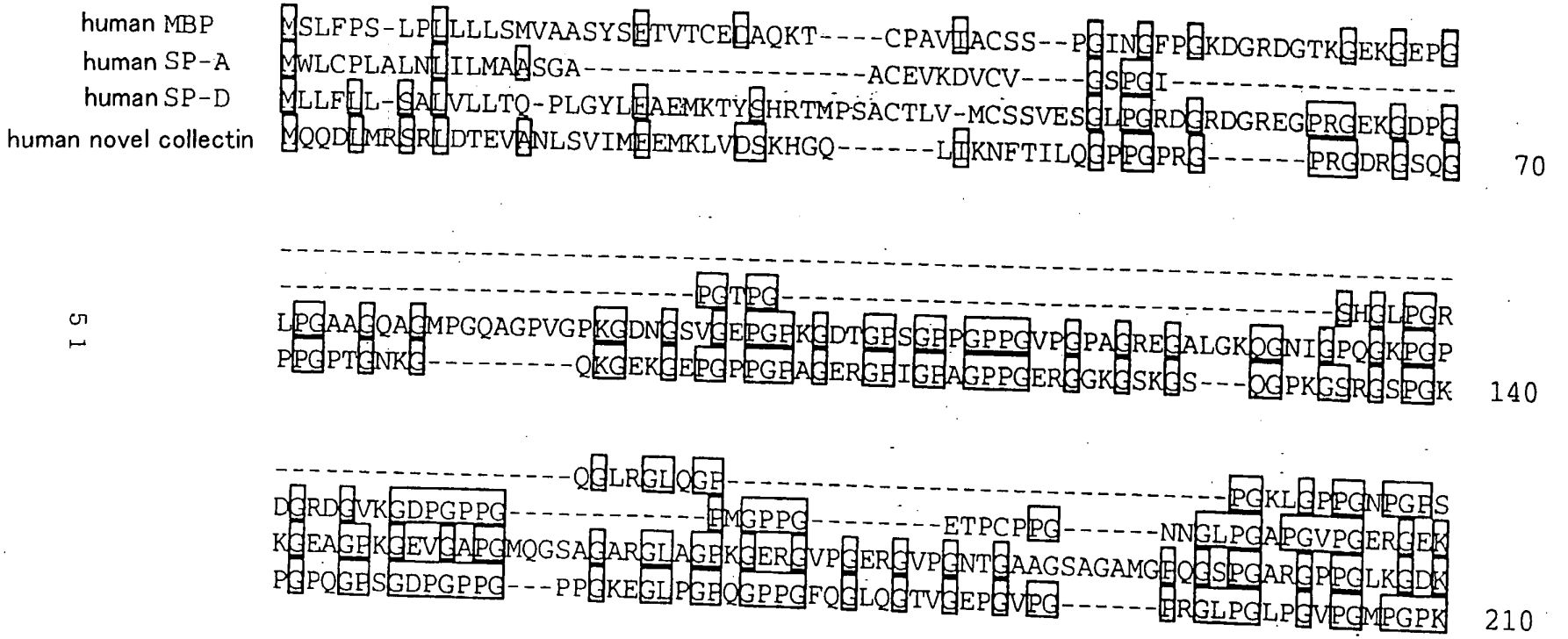

human MBP
human SP-A
human SP-D
MSLFPS-LPLLLLSMVAASYSETVTCEDAQKT----CPAVIACSS--PGINGFPGKDGRDGTKGEKGEPG
MWLCPLALTLILMA----------ASGAACEVKDVCV----------GSPG
MLLFLL-SALVLLTQ-PLGYLEAEMKTYSHRTTPSACTLV-MCSSVESGLPGRDGRDGREGPRGEKGDPG
70
IPGTPGSHGLPGRDGR----------DGVKGDPGPPGPMGPPG-------ETP----------
LPGAAGQAGMPGQAGPVGPKGDNGSVGEPGPKGDTGPSGPPGPPGVPGPAGREGPLGKQGNIGPQGKPGP
140
----------QGLRGLQGP----------PGKLGPPGNPGPSGSPGPKGQK
----------CPPGNNGLPGAPGVPGE---------RGEKGEPGERGPPGL----
KGEAGPKGEVGAPGMQGSAGARGLAGPKGERGVPGERGVPGNAGAAGSAGAMGPQGSPGARGPPGLKGDK
210

[FIGURE 2]

[FIGURE 3]

【FIGURE 4】

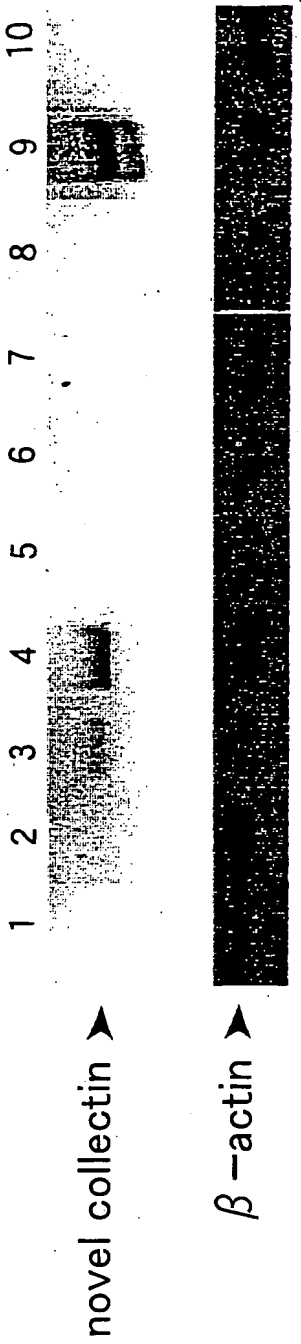

(a)
100bp
G-X-Y
BamH I
EcoR I
Sma I
Kpn I
Hind III
Sau3A I
Sau3A I
EcoR I
HPP1
HPP2
HPP3
HPP5
U
CL-P1-2-1
(b)
EcoR I
CL-P1-3-4
CL-P1-3-7
R


human MBP
human SP-A
human SP-D
human novel collectin
70
140
210

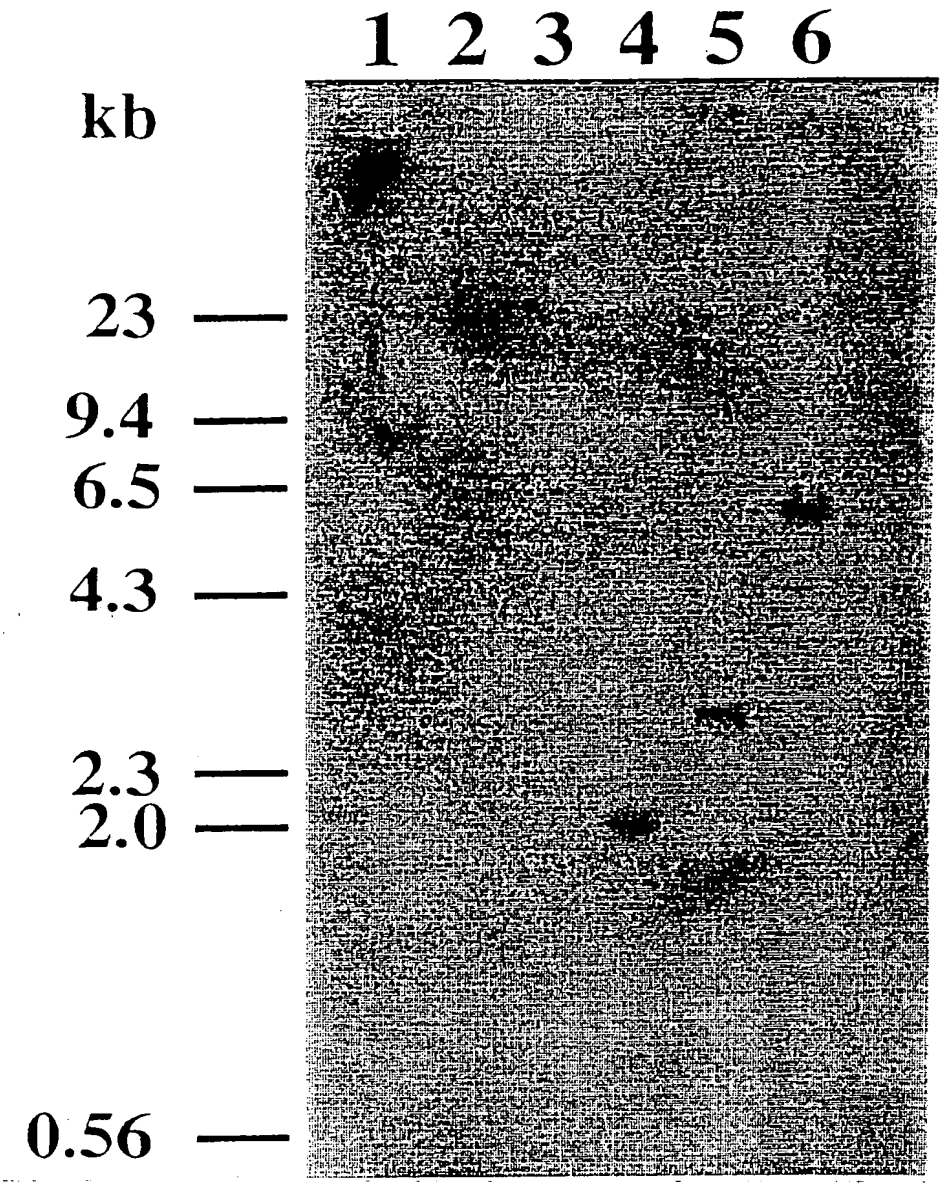
[FIGURE 5]

【FIGURE 6】

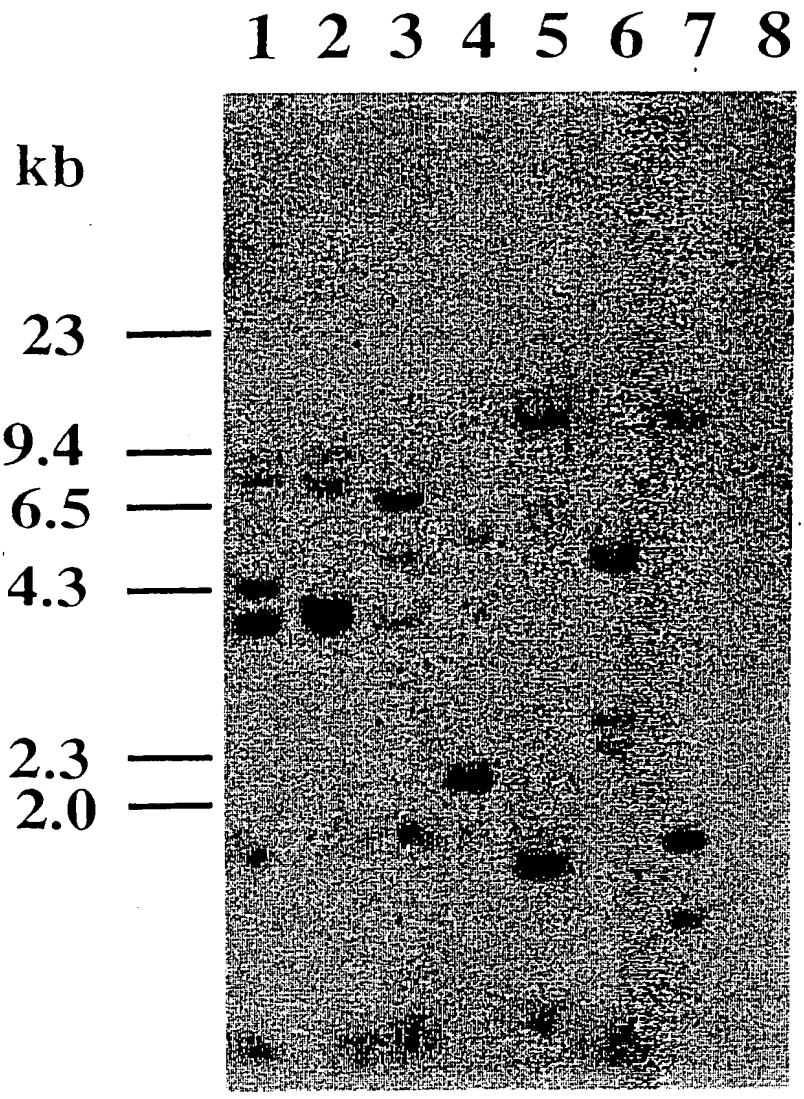
【FIGURE 7】



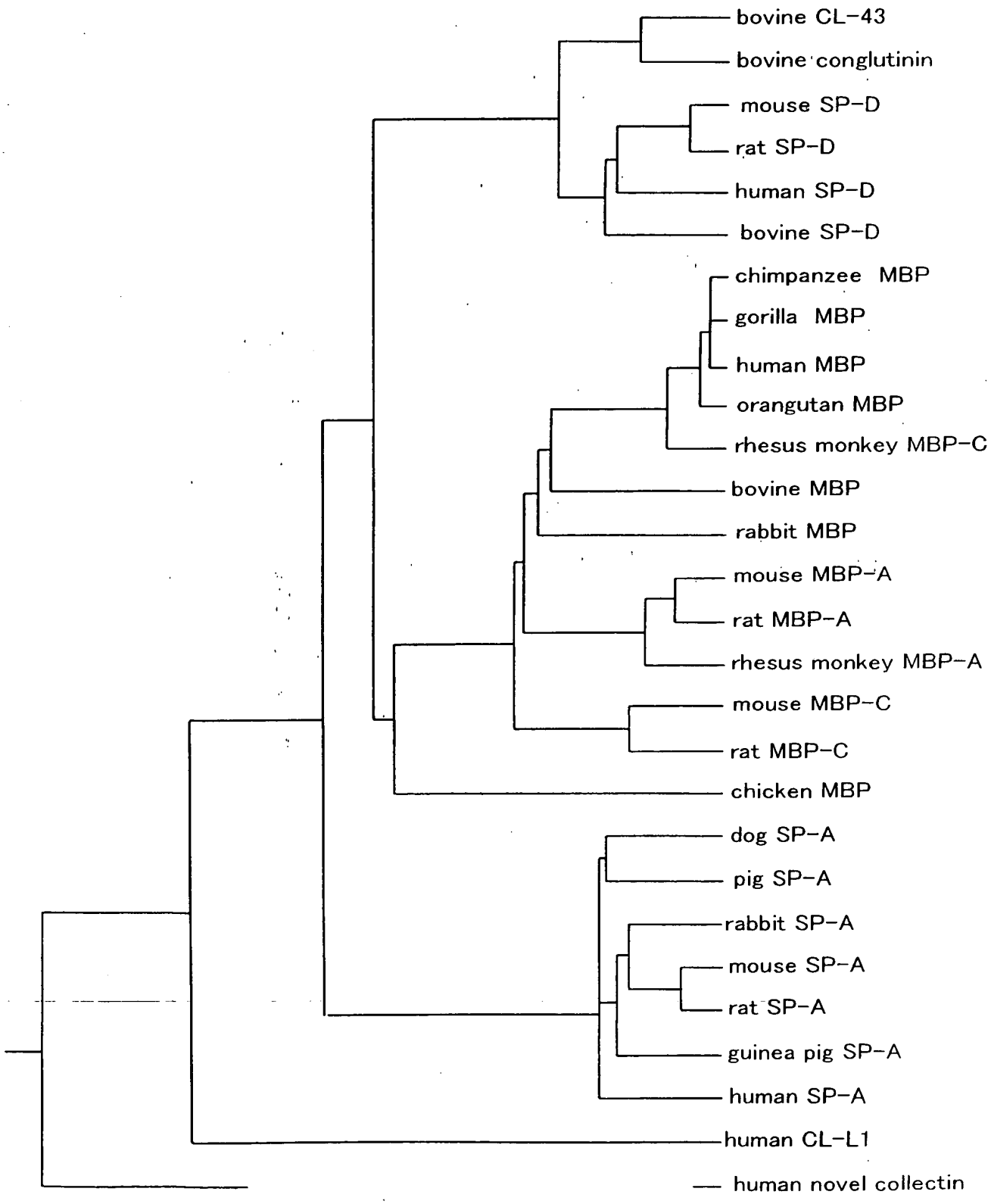
[FIGURE 8]



[FIGURE 9]



【FIGURE 10】


bovine CL-43
bovine conglutinin
mouse SP-D
rat SP-D
human SP-D
bovine SP-D
chimpanzee MBP
gorilla MBP
human MBP
orangutan MBP
rhesus monkey MBP-C
bovine MBP
rabbit MBP
mouse MBP-A
rat MBP-A
rhesus monkey MBP-A
mouse MBP-C
rat MBP-C
chicken MBP
dog SP-A
pig SP-A
rabbit SP-A
mouse SP-A
rat SP-A
guinea pig SP-A
human SP-A
human CL-L1
human novel collectin

【DOCUMENT TITLE】 ABSTRACT

【ABSTRACT】

【PURPOSES】 To provide the novel collectin which are expected to exhibit anti-bacterial, anti-viral activity or the like especially in human body.

【TECHNICAL ELEMENTS】 Collectin gene comprising a base sequence set out in SEQ ID NO: 1, and collectin protein comprising an amino acid sequence set out in SEQ ID NO: 2.

【REPRESENTATIVE FIGURE】 None